ImagingEdge.app

Tap the imaging hive-mind
Catch the opportunities

In Diagnostic Imaging Opportunities Abound…

- $40BN market by 2021
- Companies with latest tech (Hologic, Zebra): *10+% CAGR*
- Generalist companies (Siemens HC, Philips HC): *~2% CAGR*

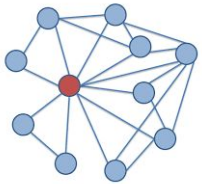
…If You Can Cut Through The Noise

- Exponentially increasing numbers of papers and abstracts
- “Technology push”
- Publications vary in quality (even at top journals / groups)

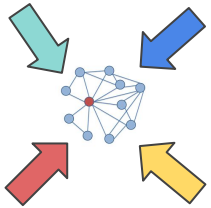
ImagingEdge detects emerging trends from diagnostic imaging research

Scrape *Labeled* Abstracts

Mine for trending phrases



Connect Trends, Labels in Graph

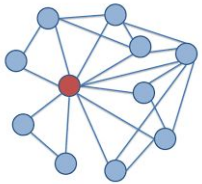


Use trained graph to mine proceedings, preprints, blogs, etc.

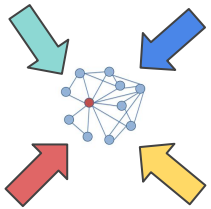
ImagingEdge detects emerging trends from diagnostic imaging research

Scrape *Labeled* Abstracts

Mine for trending phrases



Connect Trends, Labels in Graph



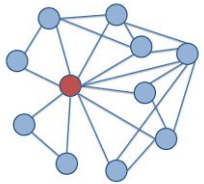
Use trained graph to mine proceedings, preprints, blogs, etc.

Goal: Catch Bleeding-Edge Trends Before They Become Publications / Products / Patents

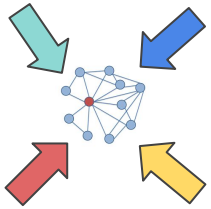
Each step requires addressing unique challenges:

How to scrape / parse / filter?

What metric to evaluate a trend?



How to connect searches with trends?



How to parse less structured sources?

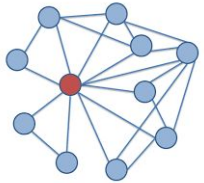
Each step requires addressing unique challenges with innovative solutions:

How to scrape / parse / filter?

Find persistent BOW ngrams using NLTK and Biopython

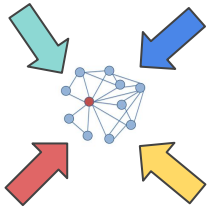
What metric to evaluate a trend?

“Trend score” combining slope & residual of linear regression across time points



How to connect searches with trends?

Build Networkx combining search terms & trending terms, derive “hotness” score



How to parse less structured sources?

Parse with BeautifulSoup, add edges to graph

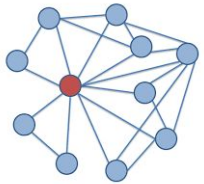
Each step requires innovative solutions with **validation challenges**:

How to scrape / parse / filter?

Are data scraping results consistent?

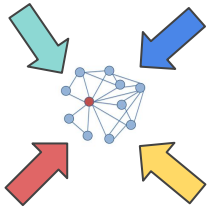
What metric to evaluate a trend?

Are trends persistent over time?



How to connect searches with trends?

Is the graph structured coherently?

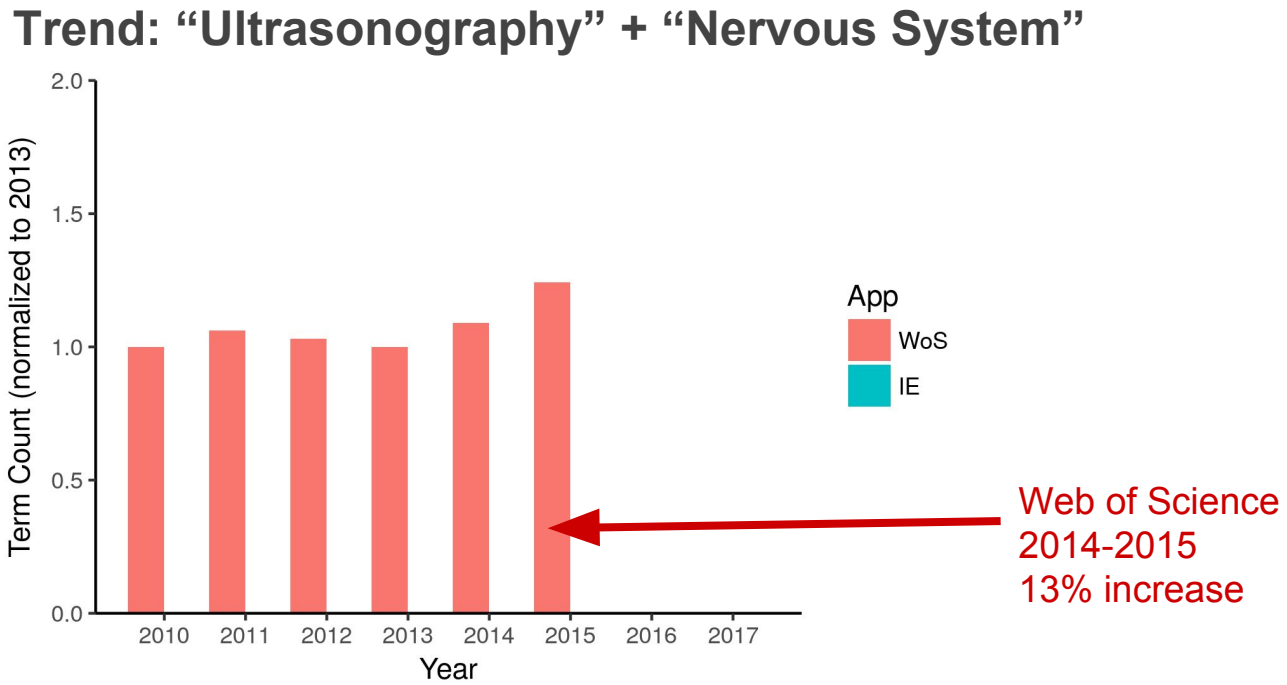


How to parse less structured sources?

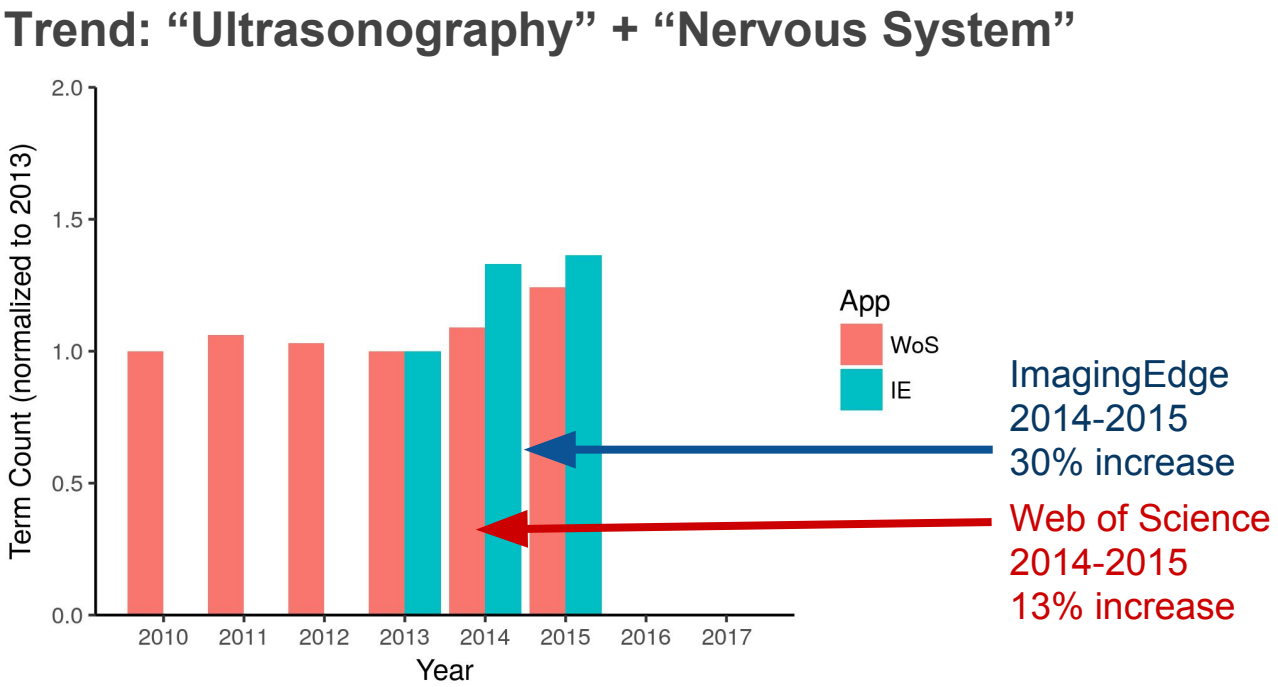
Am I correctly parsing unstructured sources?

How do I know this trend is action-able?

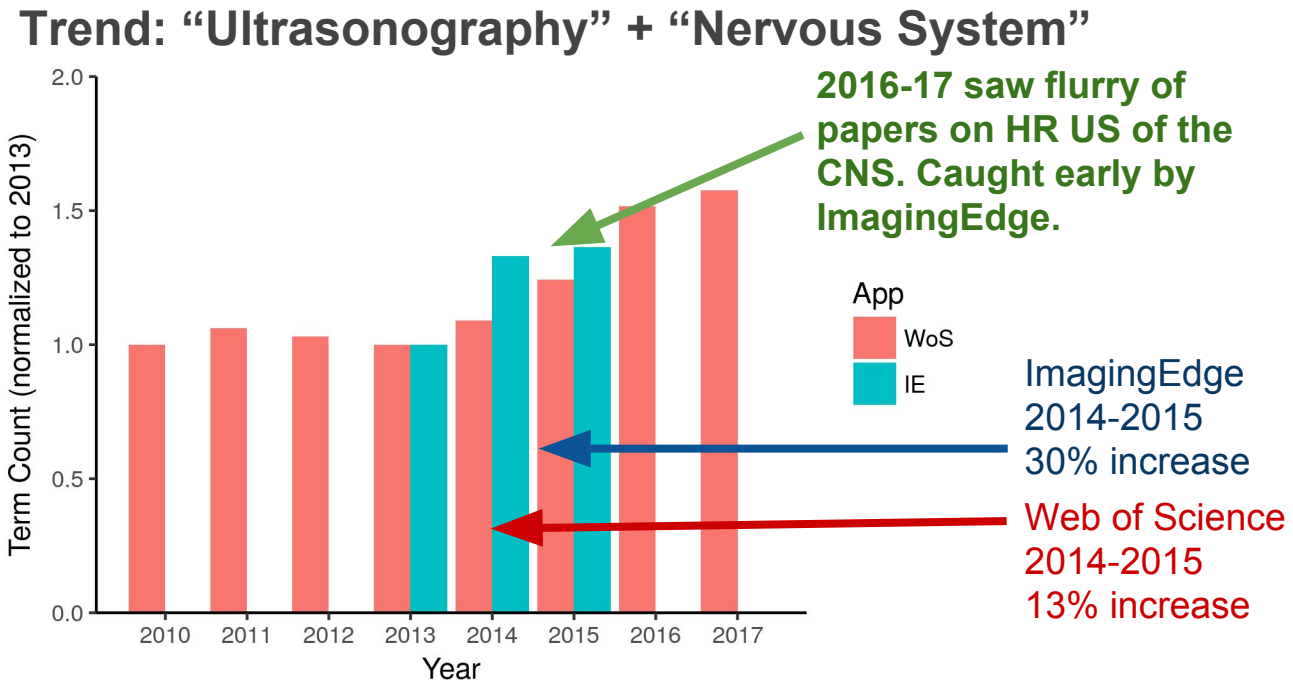
Ahead of Traditional Information Sources: Case Study



Ahead of Traditional Information Sources: Case Study



Ahead of Traditional Information Sources: Case Study



Summary: A Service To Give Companies The *Edge*

ImagingEdge:

Summary: A Service To Give Companies The *Edge*

ImagingEdge:

- Scrapes the web to organize information from diverse radiological sources

Summary: A Service To Give Companies The *Edge*

ImagingEdge:

- Scrapes the web to organize information from diverse radiological sources
- Finds trends in research activity through the noise and hype

Summary: A Service To Give Companies The *Edge*

ImagingEdge:

- Scrapes the web to organize information from diverse radiological sources
- Finds trends in research activity through the noise and hype
- Puts those trends in your hands via a simple web app

Summary: A Service To Give Companies The *Edge*

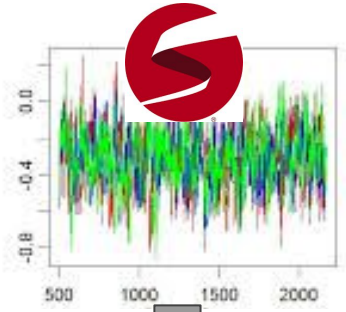
ImagingEdge:

- Scrapes the web to organize information from diverse radiological sources
- Finds trends in research activity through the noise and hype
- Puts those trends in your hands via a simple web app
- Tips you off to the next big imaging plays before they happen

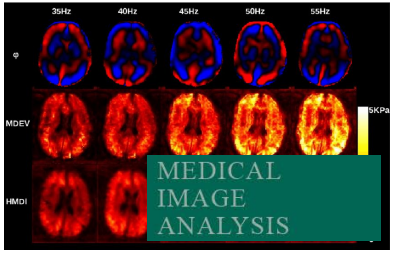
About Eric Barnhill

Juilliard DANCE DRAMA MUSIC

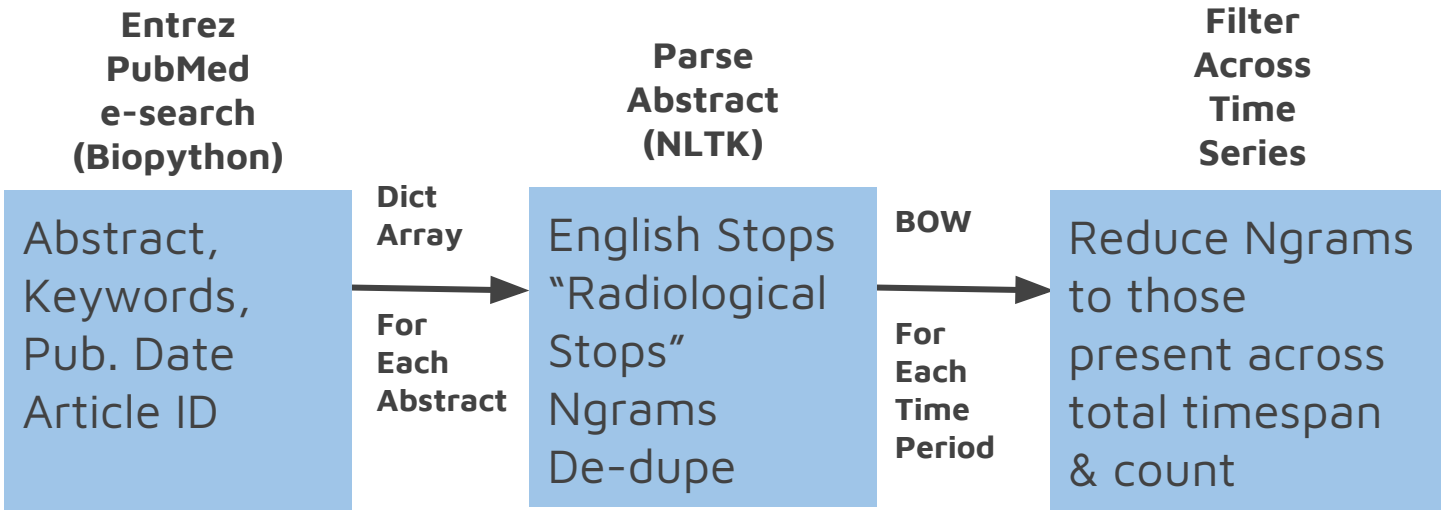
Dalcroze Society of America



CHARITÉ UNIVERSITÄTSMEDIZIN BERLIN

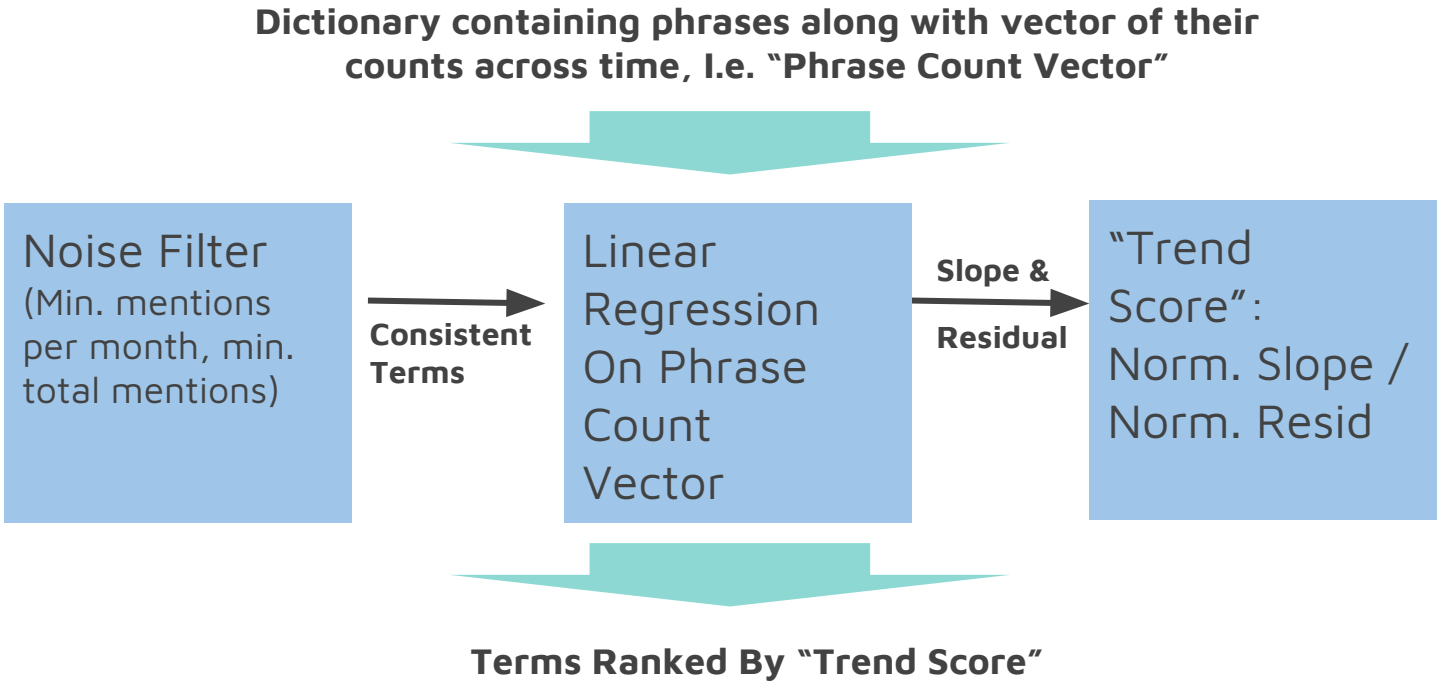


ImagingEdge methodologies: Scraping PubMed

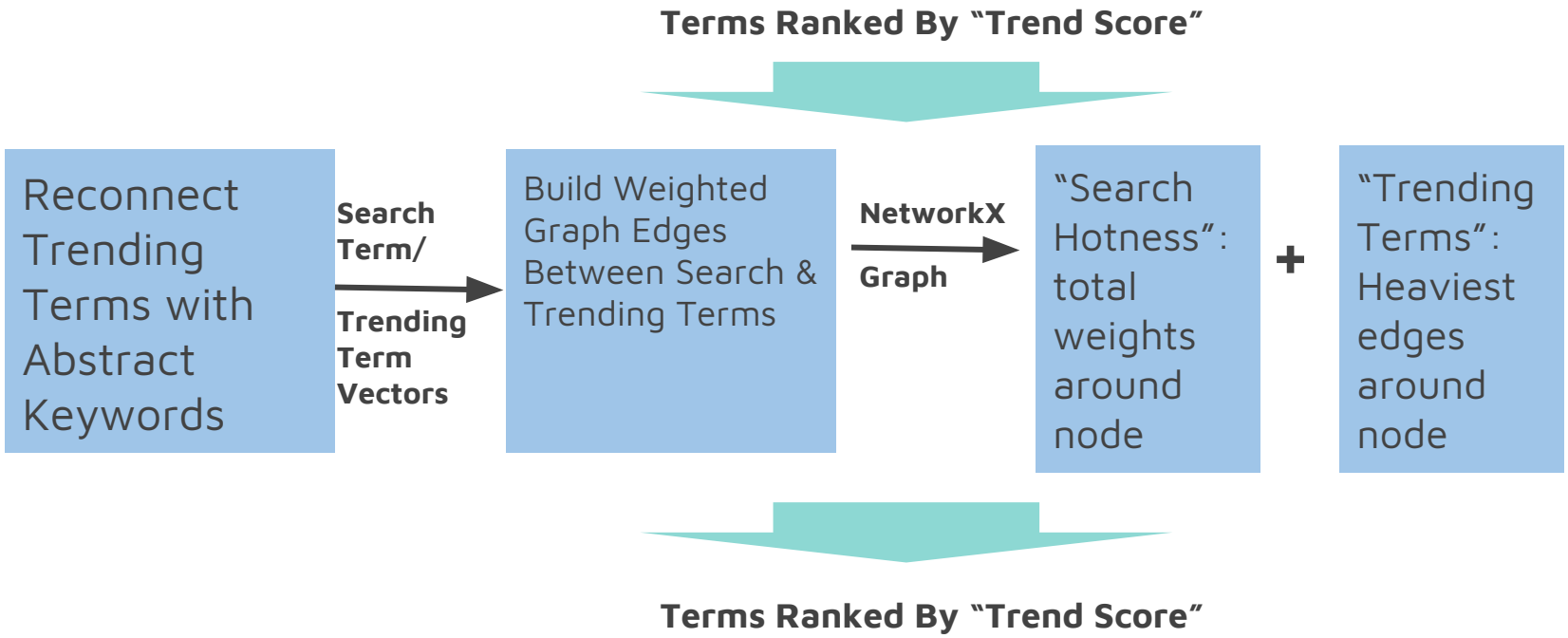


Dictionary containing phrases along with vector of their counts across time

ImagingEdge methodologies: Detecting Trends



ImagingEdge methodologies: Building the Imaging Edge Graph

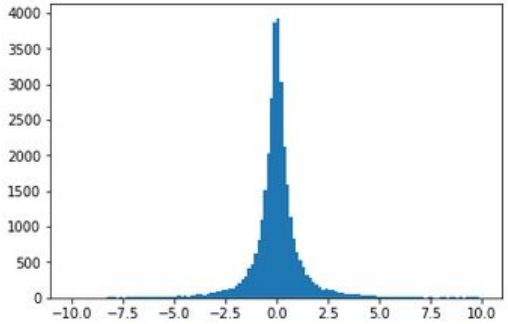


Data Exploration and Validation: Statistical Properties of Scraped Data

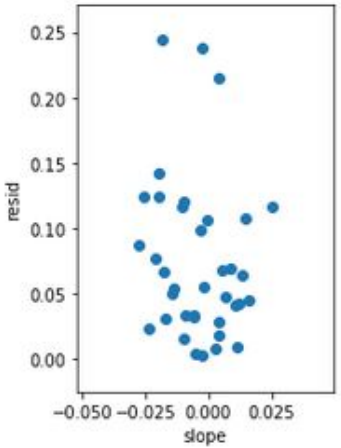
Trend linear fit slopes are centered and normal

[coming soon]

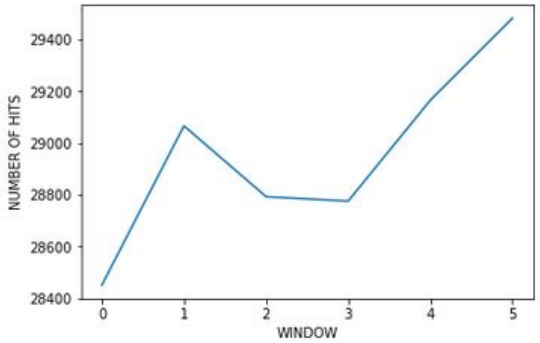
"Trend score" distributions robust, centered and skew-free



Residuals independent of trends

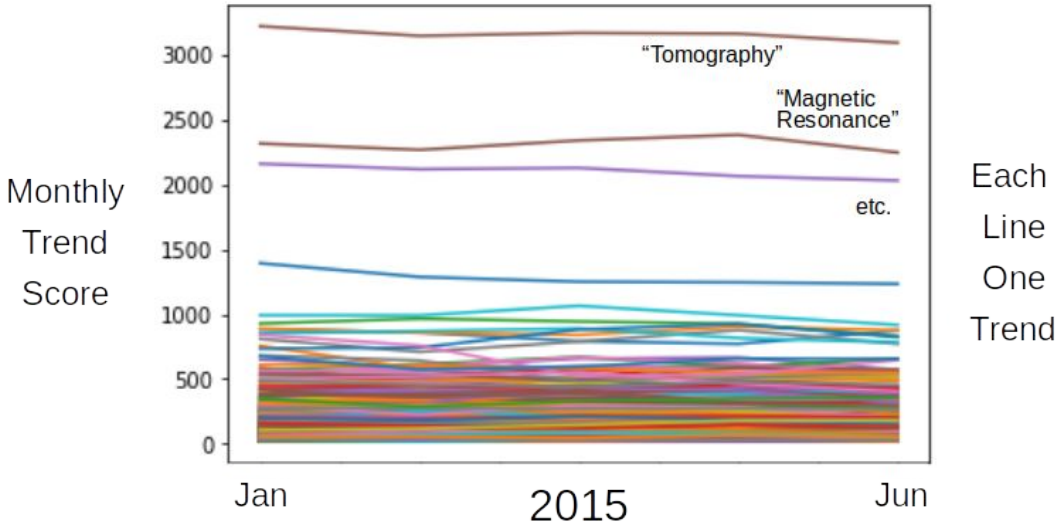


Scrapes are consistent across windows



Data Exploration and Validation: Trends Show Robustness Over Time

Time-series trends are persistent and stable across rolling time windows



Likelihood ratio test validation:

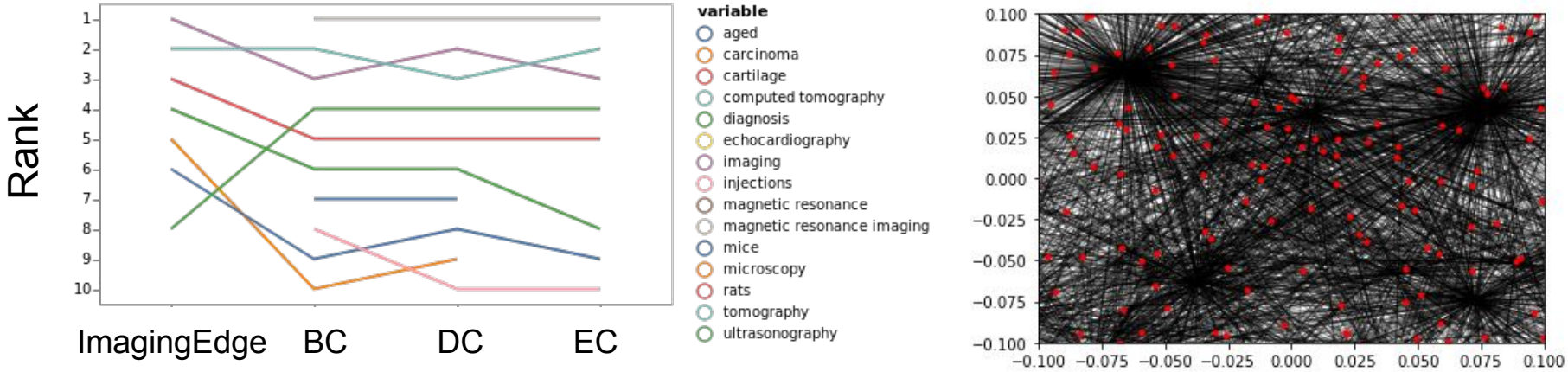
y ~ ß0 vs. y ~ (1 | ß0)

AIC delta: -464 BIC delta:- 457
(-10 considered good)
Random intercept model better fit

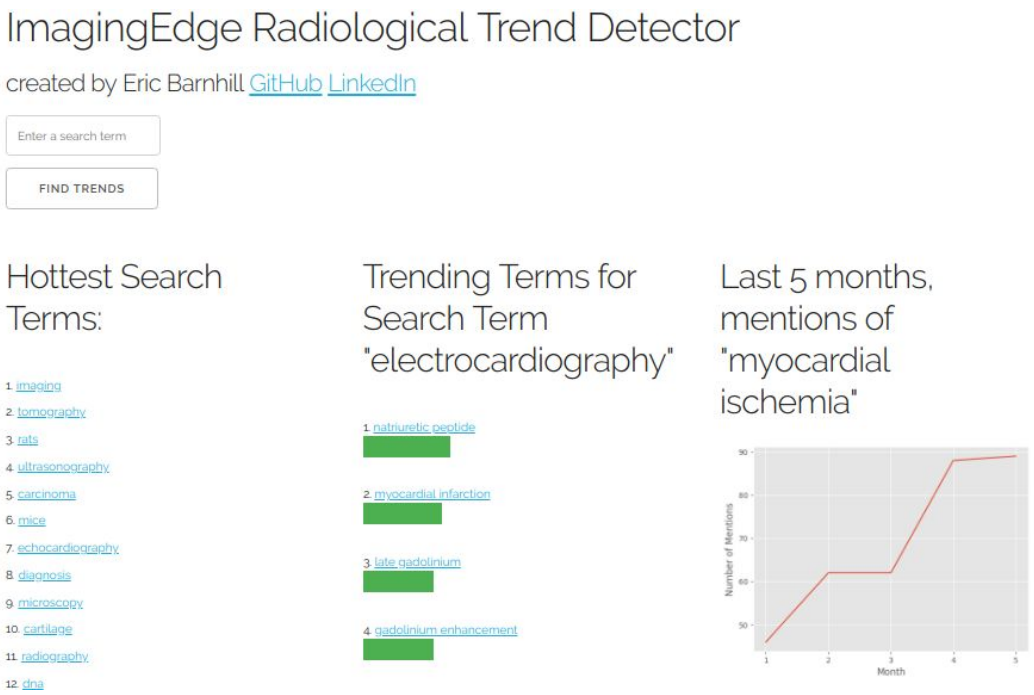
Additional "null test" (same model, rend scores scrambled)
AIC delta: +6 BIC delta:+13
Single-intercept model better fit

Data Exploration and Validation: Graph influence measures coincide with "hotness"

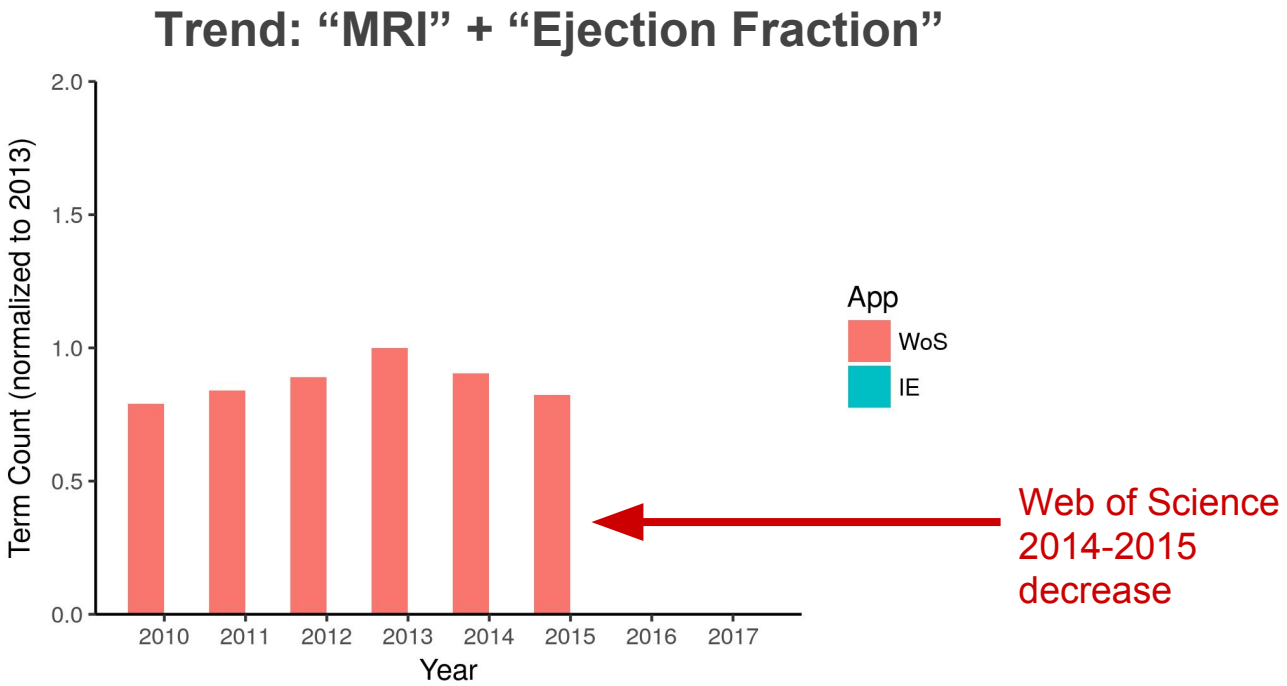
In graph, "hot search terms" dominate all centrality measures (though less with between centrality)



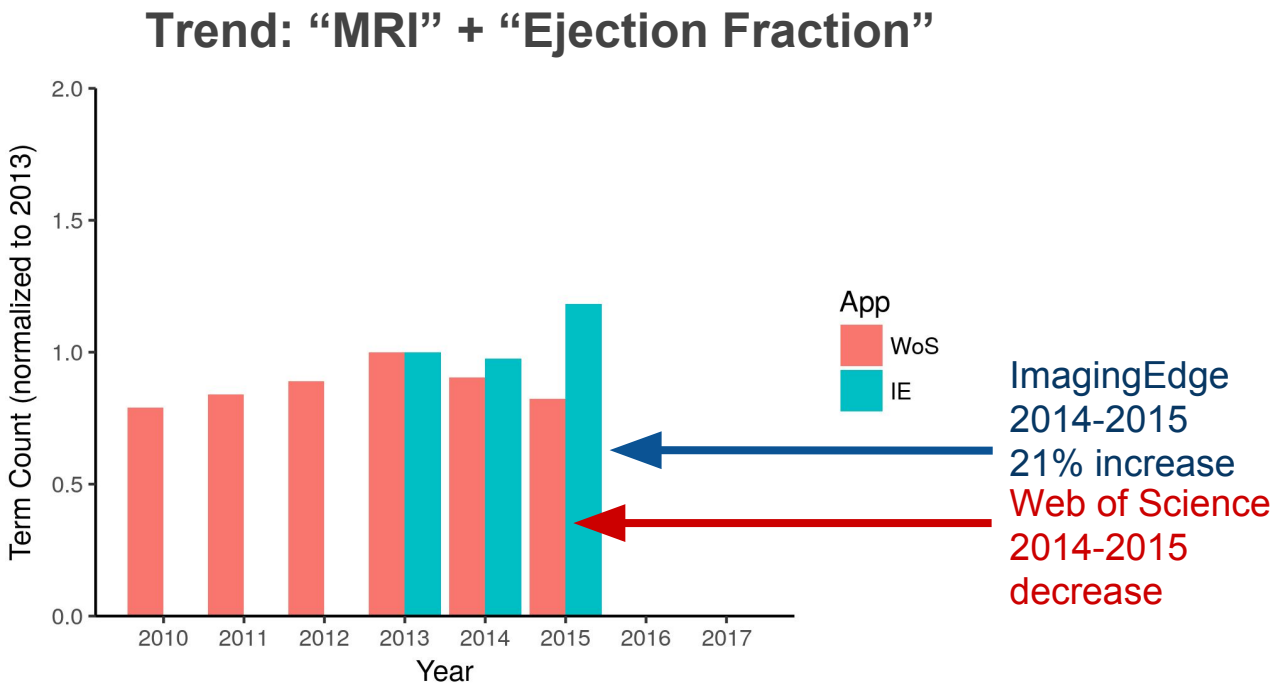
ImagingEdge.app in action



Ahead of Traditional Information Sources: Case Study



Ahead of Traditional Information Sources: Case Study



Ahead of Traditional Information Sources: Case Study