Northwestern | SCHOOL OF COMMUNICATION

Computational Journalism Lab

AI-Driven Lead Recommenders for Science Journalists

Sachita Nishal and Nick Diakopoulos

Modern challenges in science journalism

The evolution of the science journalist

Modern challenges in science journalism

The evolution of the science journalist

Changing attitudes to science

A shift from “gee-whiz” reporting and “fan-boy coverage” in the 60s/70s to “**independent inquiry** into the scientific enterprise and the illumination of research with all its **wonderfully complex human interactions**” (Blum, 2021)

Modern challenges in science journalism

The evolution of the science journalist

Changing attitudes to science

A shift from “gee-whiz” reporting and “fan-boy coverage” in the 60s/70s to “**independent inquiry** into the scientific enterprise and the illumination of research with all its **wonderfully complex human interactions**” (Blum, 2021)

New responsibilities

Science journalists have also taken on **a variety of new roles**, such as those of the **curator, the civic educator, the public intellectual, the watchdog,** etc. (Fahy and Nisbet, 2011)

Modern challenges in science journalism

A changing media landscape in the digital age

Modern challenges in science journalism

A changing media landscape in the digital age

Novel modes of engagement with sources and audiences

Web-savvy journalists have the opportunity to **tap into new informational networks** and **facilitate greater dialogue and transparency** about science via the internet (Allan, 2011)

Modern challenges in science journalism

A changing media landscape in the digital age

Novel modes of engagement with sources and audiences

Web-savvy journalists have the opportunity to **tap into new informational networks** and **facilitate greater dialogue and transparency** about science via the internet (Allan, 2011)

… but also more noise on the internet that impacts reporting

“A dizzying array of **science information, misinformation, and commentary** that can be hard to sort through” (Russell, 2009)

Modern challenges in science journalism

And the exploding volume of scientific output itself!

Modern challenges in science journalism

And the exploding volume of scientific output itself!

There's a LOT of research being published out there

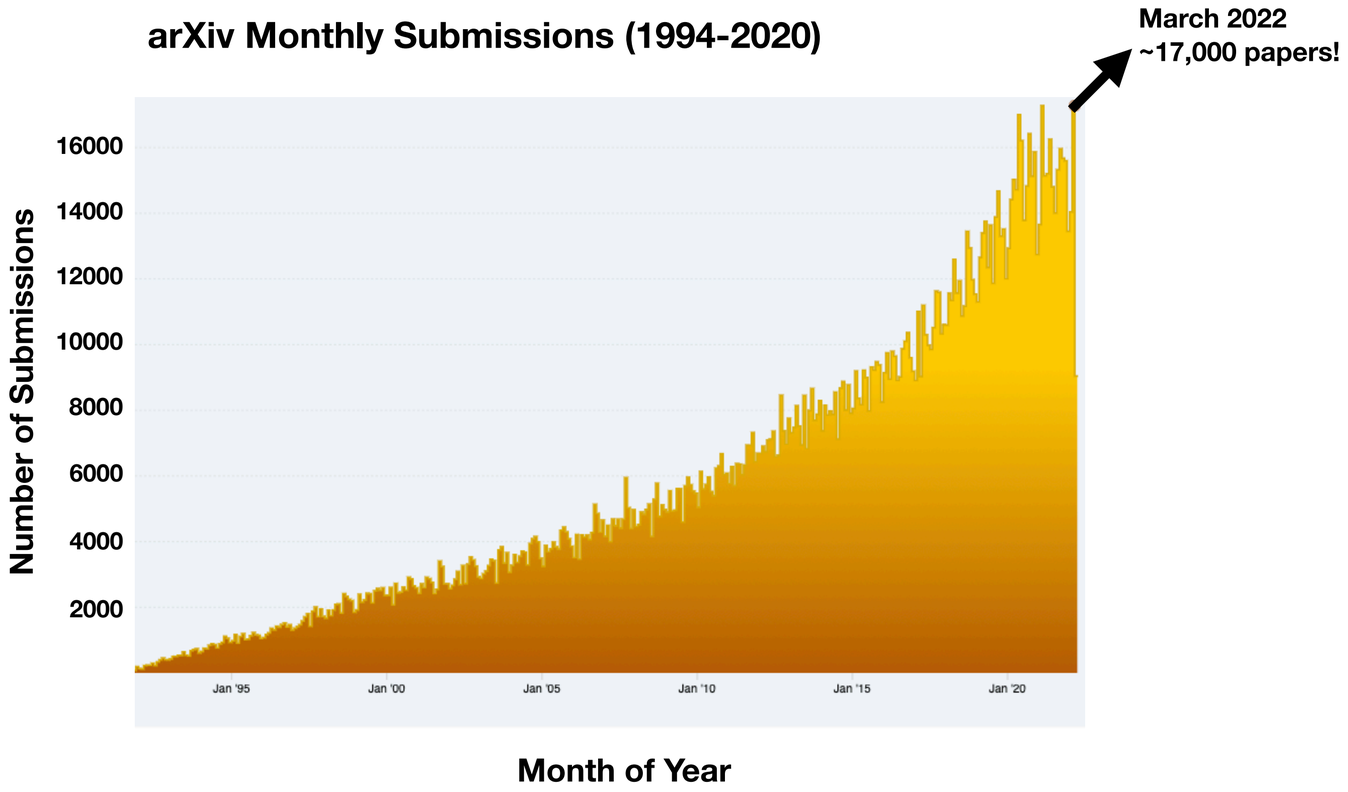
> "Between August 2019 and August 2020, Scopus added **3.29 million new records**"
> (Scopus Blog, 2021)

Modern challenges in science journalism

And the exploding volume of scientific output itself!

There's a LOT of research being published out there

"Between August 2019 and August 2020, Scopus added **3.29 million new records**"
(Scopus Blog, 2021)



How can we support science journalists as they grapple with this firehose of information from various sources, sift through it for relevant leads, and try to kick-start the news process?

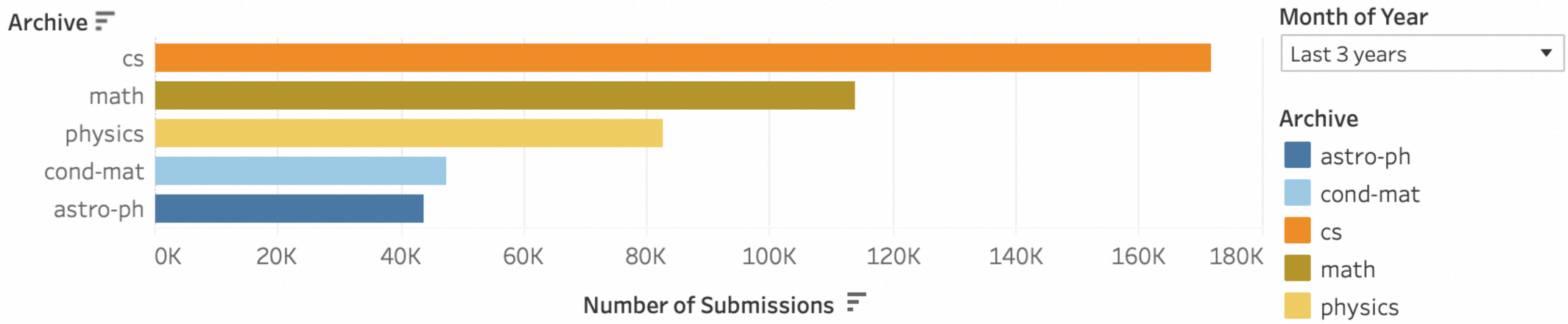
Computational News Discovery

“The **use of algorithms** to **orient editorial attention** to **potentially newsworthy events** or information prior to publication” (Diakopoulos, 2020)

Computational News Discovery

"The **use of algorithms** to **orient editorial attention** to **potentially newsworthy events** or information prior to publication"

Scope of our work: CND for **arXiv pre-prints** in the **Computer Science category**, and with more newsworthy sub-categories that we selected.



Building automatic lead recommenders

Building automatic lead recommenders

Design Goals

- Provide time and information subsidies
- Personalize to journalistic interest
- Ensure journalistic autonomy

Building automatic lead recommenders

Design Goals

Provide time and information subsidies

Personalize to journalistic interest

Ensure journalistic autonomy

Implementation

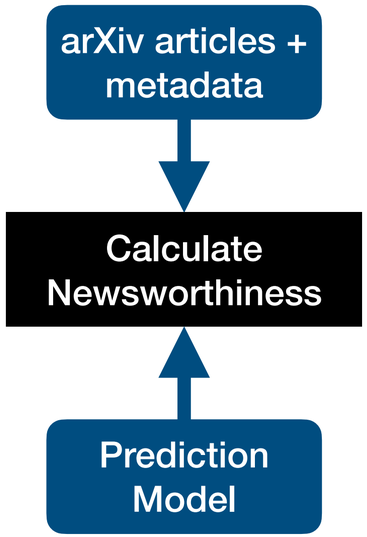
Develop ML-driven pipeline to **suggest newsworthy research** that is **relevant to specific publications** that journalists would like to write stories for.

Contextualize potential leads, and provide a **news angles for creative story ideas**

Deploy the recommender with professional journalists and **collect feedback** for design iterations

Lead recommender pipeline

Lead recommender pipeline



Measuring newsworthiness

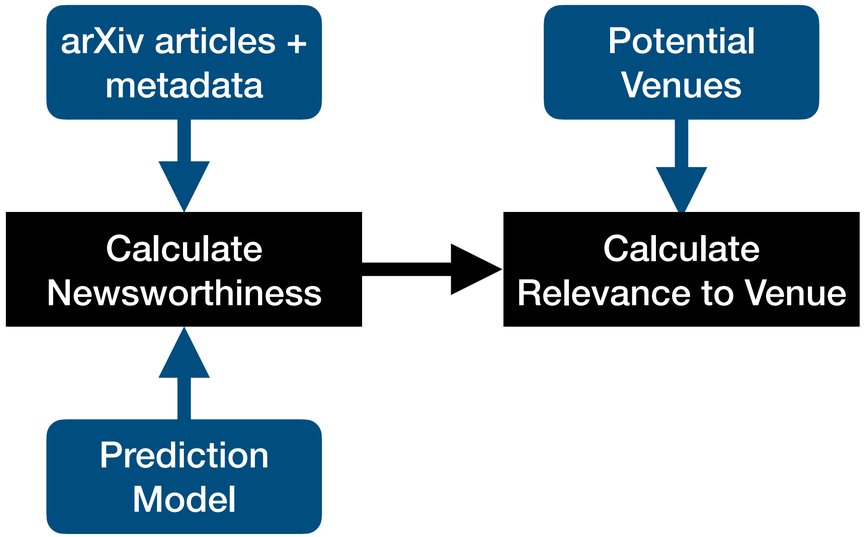
In our previous work (Nishal and Diakopoulos, 2022), we have demonstrated that non-expert crowd-workers can recognise certain **news values in scientific abstracts**, and that these **correlate moderately well with expert journalists' judgments**.

- Actuality
- Controversy
- Impact Magnitude
- Impact Valence

We aggregated these news values to obtain a **crowd newsworthiness score,** and trained a Random Forest model to **predict it using the text of scientific article**. Prec@K=10: 80%

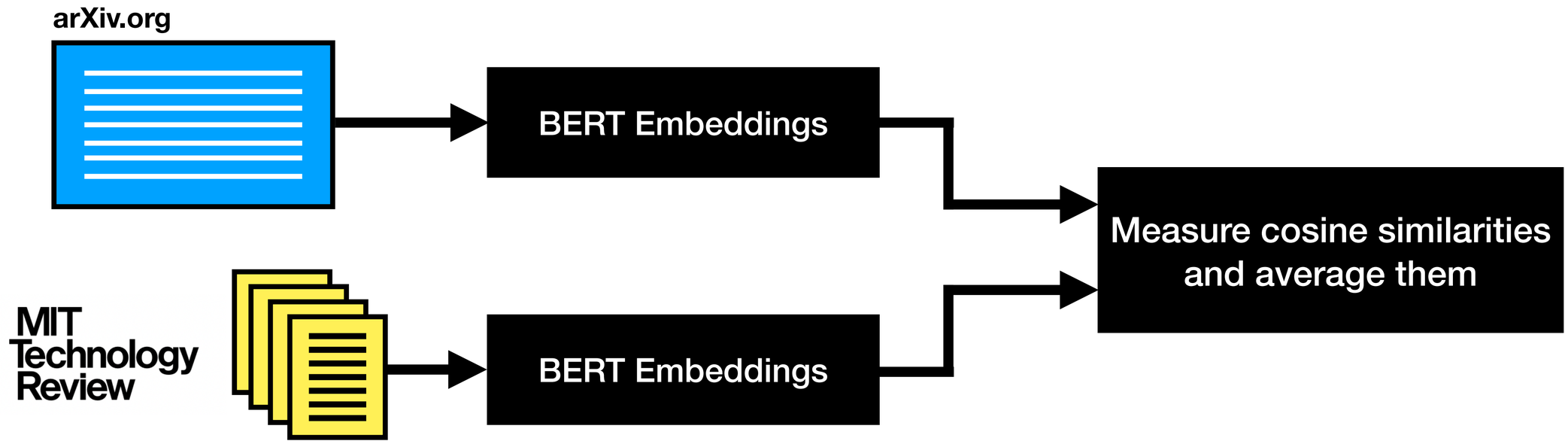
Nishal S. and Diakopoulos N. 2022. From Crowd Ratings to Predictive Models of Newsworthiness to Support Science Journalism. In Proceedings of the ACM on Human-Computer Interaction. CSCW (2022) **(forthcoming)**

Lead recommender pipeline

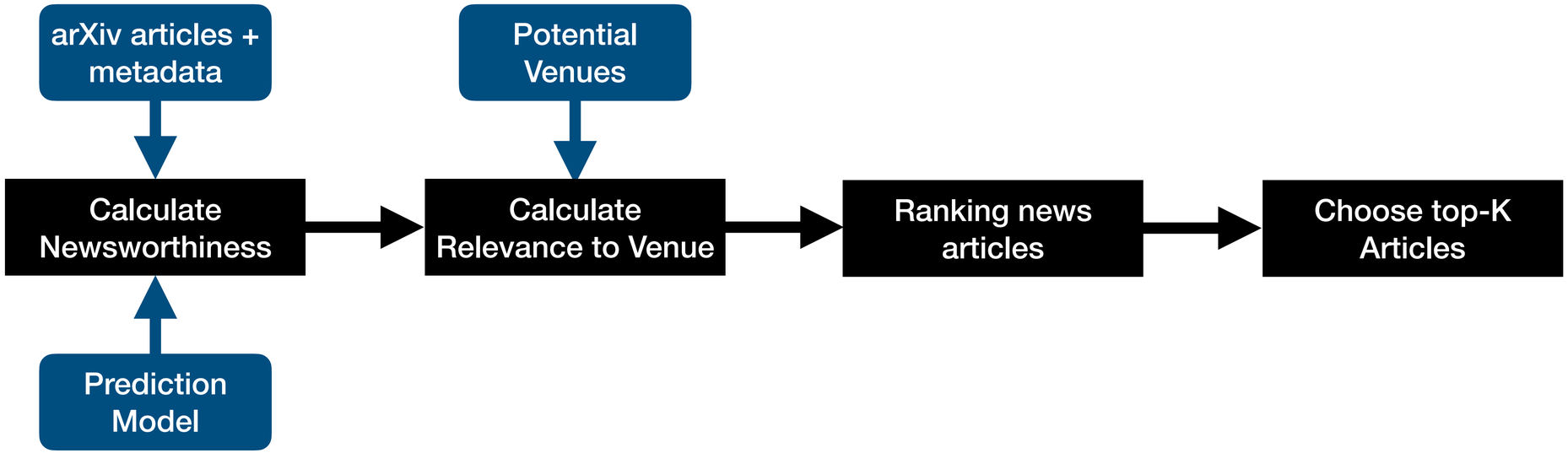


Personalizing to publication venues

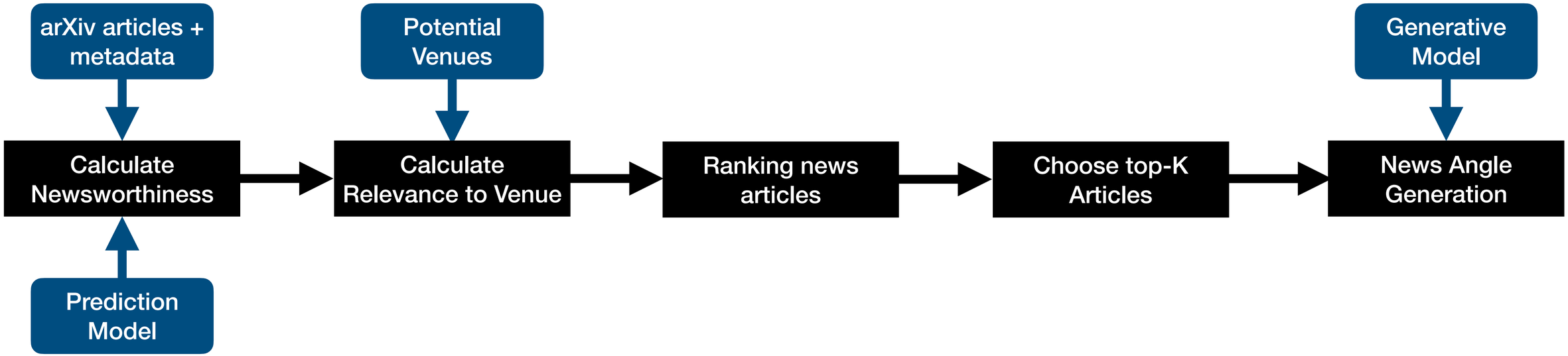
We collect **historical data from thirty news outlets** that cover stories in science and technology + measure the **similarity of a given research abstract** to their published stories.



Lead recommender pipeline



Lead recommender pipeline

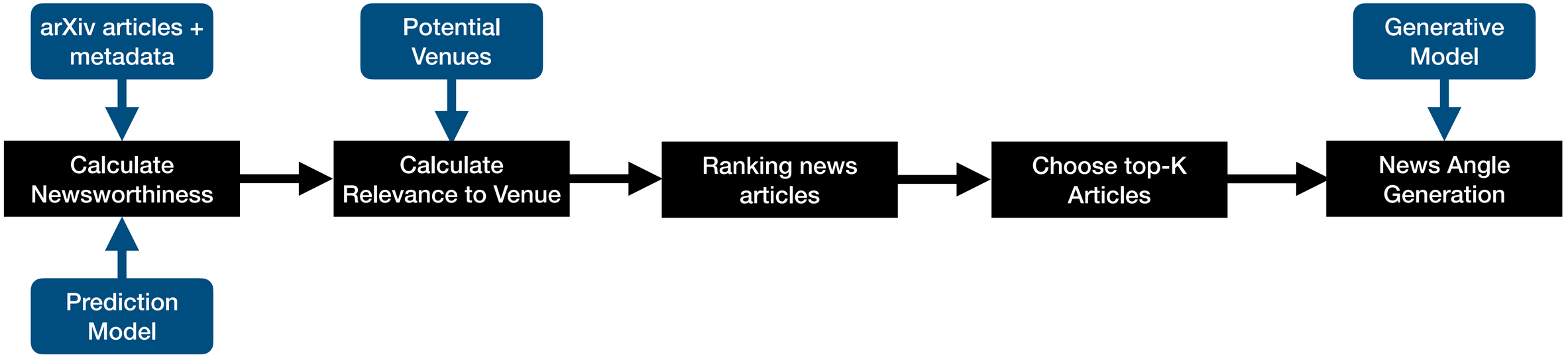


News angle generation

Fine-tune a pre-trained GPT-3 model to **generate news angles** based on the **title and the abstract of a given article**

Assess predictions on the basis of **automated metrics** and the judgements of **human evaluators**

Lead recommender pipeline



Deploy with professional journalists
+
conduct semi-structured interviews

Deployment studies

Recruit science and technology journalists, and **collect feedback** about the recommender system using deployments and semi-structured interviews.

User-interface is currently under development, and we are **actively looking for journalists to participate** in our study!

If you or anyone you know might be interested, **we'd love to hear from you**!

Get in touch!

Sachita Nishal
PhD Student
School of Communication
Northwestern University

Email: nishal@u.northwestern.edu
Twitter: @nishalsach

Nick Diakopoulos
Associate Professor
School of Communication
Northwestern University

Email: nad@northwestern.edu
Twitter: @ndiakopoulos