Computational News Discovery:
Designing Journalistic Tools to Support Algorithmic Lead Generation

MediaFutures Annual Meeting
September 29, 2021

Nick Diakopoulos, Northwestern University
nad@northwestern.edu | @ndiakopoulos

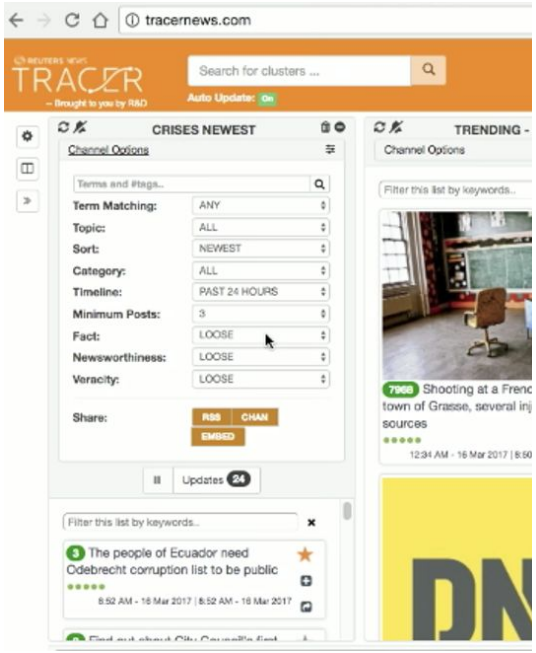
146,780

News stories published by just a small sample of 87 media sources in the US over 30 days from Aug 23 - Sept 23

Source: Media Cloud, US Top Sources 2018

Computational News Discovery (CND)

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



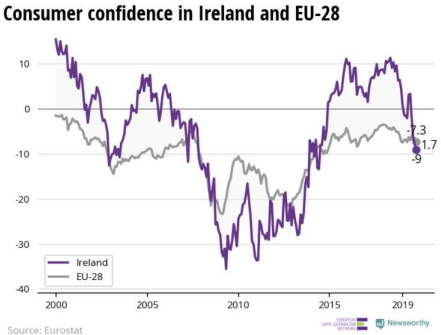
THE CONFIDENCE OF CONSUMERS IS DECREASING MORE RAPIDLY IN IRELAND THAN IN THE EU

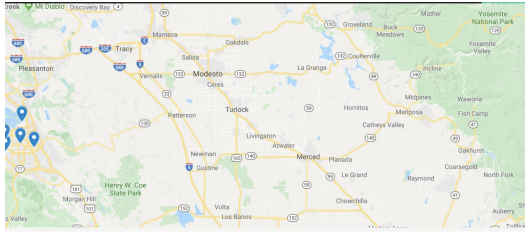
In October the confidence of consumers in Ireland decreased significantly compared to the same month last year, down from 6.4 to -9.

In the same time, EU countries consumers confidence have decreased somewhat, down from -5 to -7.3 on average.

Ireland now has the seventh lowest confidence of consumers in the EU compared to the other 28 countries. Lithuania has the highest confidence of consumers (3.8 points), while Bulgaria ranks last (-22 points).

Read more





Welcome to Agenda Watch

Learn about newsworthy public meetings before they happen. Track government action at the city, county and district levels. Find local stories that matter to readers.

Agenda Watch pulls in meeting agendas from around the U.S. and reads them so you don't have to. Our software sifts through documents, then sends you an email digest of the agenda items relevant to your region and beat.

How to Design CND Systems?

Interviewed 18 people in 2018-19 who use or build these tools.

Tools like RADAR, Newsworthy, Klaxon, FullFact, Chequado, as well as other internal projects.

Twenty questions about tool use in newswork, utility, volume, fit into workflow, trustworthiness, ethical issues.

Qualitative thematic analysis of interview transcripts

The Newsroom Attention Economy

- Reporter Willingness
- External Contextual Factors
- Monitoring Scope
- UI Design

Newsworthiness Assessment

Quality Assurance

Computational News Discovery: Towards Design Considerations for Editorial Orientation Algorithms in Journalism. Digital Journalism. 2020.

WaPo Lead Locator

Washington Post Tool build to help data mine the US voter file for 2020 Election.

Goal to augment journalists' ability to find interesting demographic patterns in US jurisdictions that they would then visit to report about. **Where** to go? **Who** to talk to?

Newsworthiness: **Novelty** and contrast to central tendency (deviance); **Political Relevance** (party margins and flips); **Magnitude** (county pop)

UI Design: search and browse, location context

Generating Location-Based News Leads for National Politics Reporting. Proc. Computation + Journalism Symposium. 2020.

Pennsylvania: White Non-College-Educated Turnout

Politics

In the key state of Pennsylvania, vote counting could stretch beyond election night

October 12, 2020 | 10:01 AM EDT

In Erie County, Pa. — a swing county in a battleground state — voters and election officials are trying to navigate the challenges of holding an election in the midst of a pandemic.

Chester County, Pennsylvania

Noteworthy Points

- The White Non-College-Educated Turnout Rate is 73.4%, the highest in the state. The state average

Political Relevance

- From 2012 to 2016 the county shifted left by 9.6% (flipping from R to D). And from 2016 to 2018 the

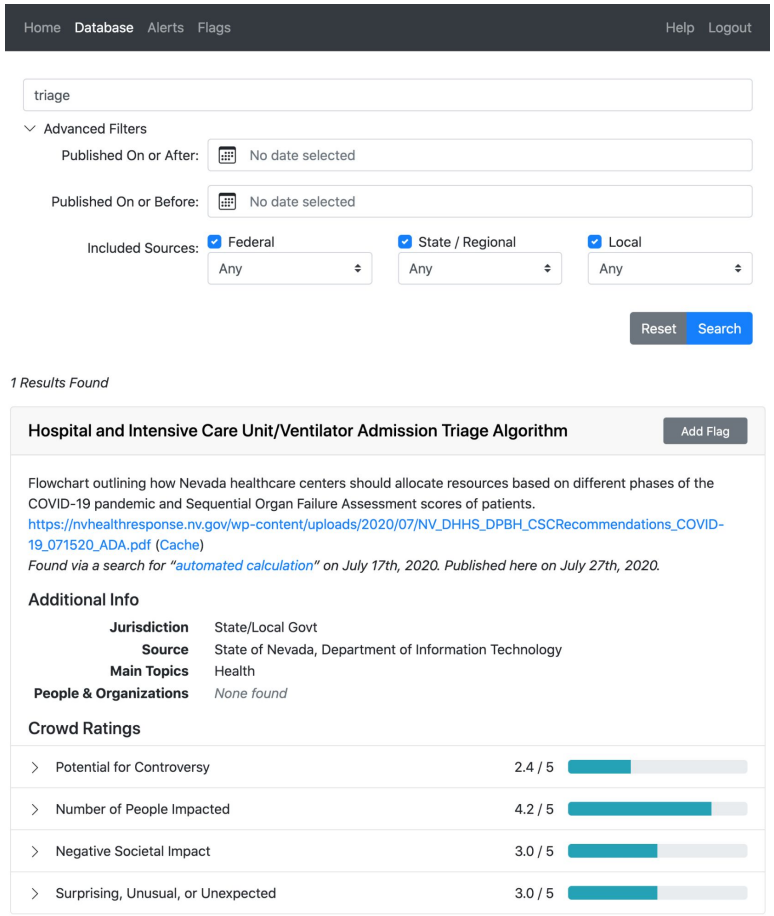
Algorithm Tips

System to enable algorithmic accountability reporting and support news discovery of Algorithmic Decision Making (ADM) tools in U.S. government. **What's** going on with ADMs in government?

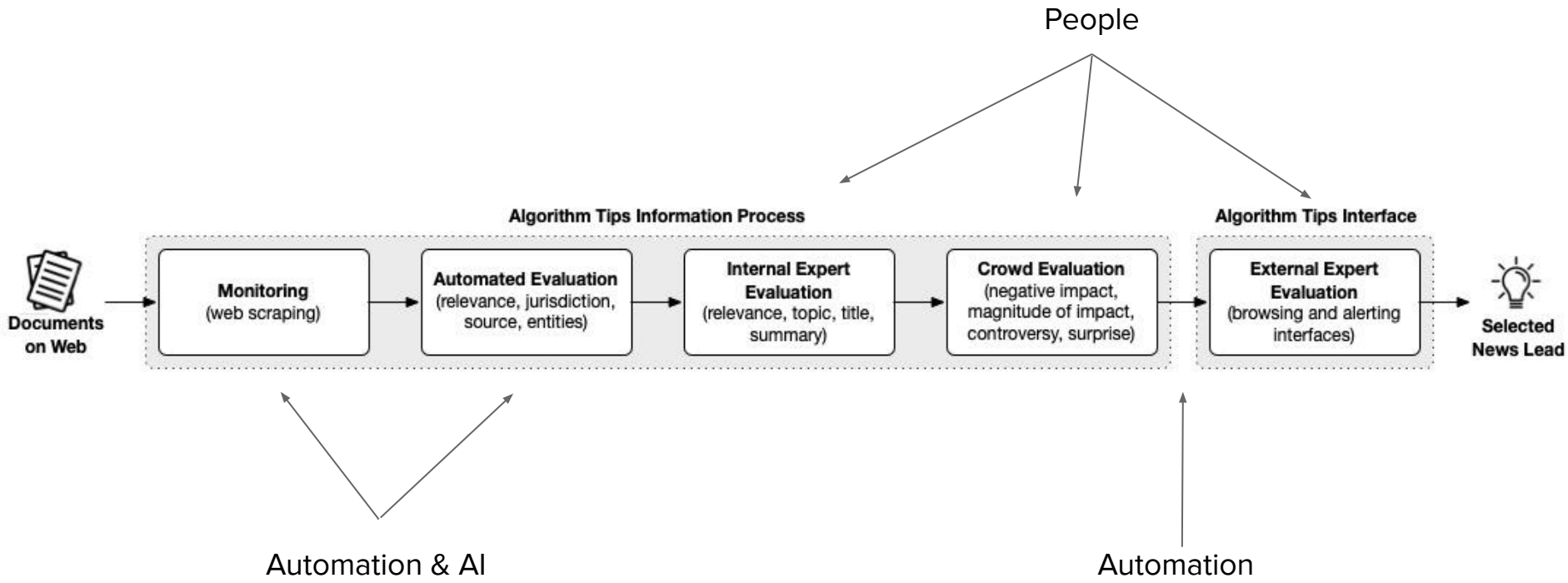
https://db.algorithmtips.org/

Design Goals

- Be sensitive to human effort and attention
- Enable initial verification and quality assessment
- Enable newsworthiness evaluation



Algorithm Tips: A Hybrid Process



Algorithm Tips Evaluation

Longer deployment: 8 weeks with 8 professional journalists with experience covering algorithms in society.

Interview study asking about features, relevance, newsworthiness, crowd ratings, usage, time spent, fit with broader workflow, ethical issues

Qualitative thematic analysis of interview transcripts

Design goals were met and well-received, but some avenues for improvement

Contextual variance: finding the right parameters to allow it to suit different needs. Definitions.

Newsworthiness judgements: could be tailored even more to the domain. Crowdsourced ratings has mixed utility.

Sustainability and effort: new work, ongoing periodic work, but we *need* people in the process to ensure high quality. Media Management questions.

Towards Understanding and Supporting Journalistic Practices Using Semi-Automated News Discovery Tools. Proc. ACM (CSCW). 2021.

Read More...

N. Diakopoulos. Computational News Discovery: Towards Design Considerations for Editorial Orientation Algorithms in Journalism. *Digital Journalism*. 2020.

N. Diakopoulos, M. Dong, L. Bronner, J. Bowers. Generating Location-Based News Leads for National Politics Reporting. *Proc. Computation + Journalism Symposium*. 2020.

D. Trielli, J. Stark and N. Diakopoulos. Algorithm Tips: A Resource for Algorithmic Accountability in Government. *Computation + Journalism Symposium*. 2017.

N Diakopoulos, D. Trielli, G Lee. Towards Understanding and Supporting Journalistic Practices Using Semi-Automated News Discovery Tools. *Proc. ACM (CSCW)*. 2021.

Thanks! Questions?

Contact:
Nick Diakopoulos, nad@northwestern.edu, @ndiakopoulos
http://www.nickdiakopoulos.com/

In Collaboration:
Lenny Bronner, Madison Dong, Jeremy Bowers, Grace Lee, Daniel Trielli