

Sim**PPL**



Rebuild Digital Trust

team@simppl.org

All rights reserved. © SimPPL

2024.

About

- ❖ SimPPL is a global community of 200+ engineers and researchers committed to improving **digital trust**.
- ❖ We are a nonprofit that empowers **young innovators** to become leaders.
- ❖ We build responsible computing tools to make **reliable information** accessible for the **global majority**.

Research

Information systems research informs newsrooms, nonprofits, fact-checkers, and inter-governmental agencies.

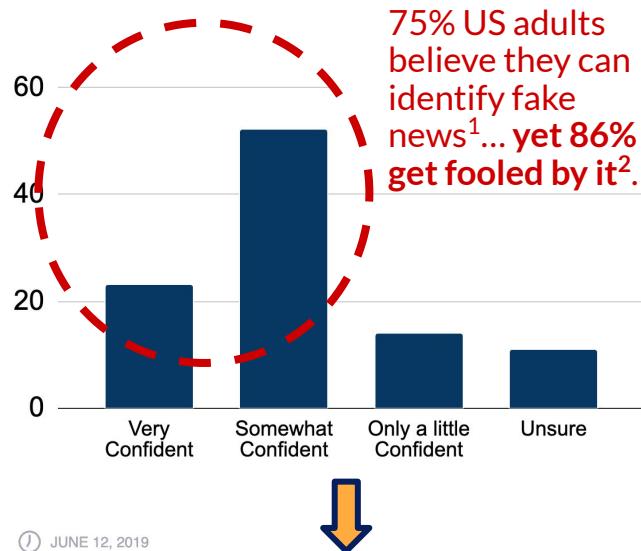
Products

Build scalable systems to advance digital trust.
Deploy sustainable technology for social good
with global partners.

Supporting UN Sustainable Development Goals



Online Trust Declining: The Misinformation Effect

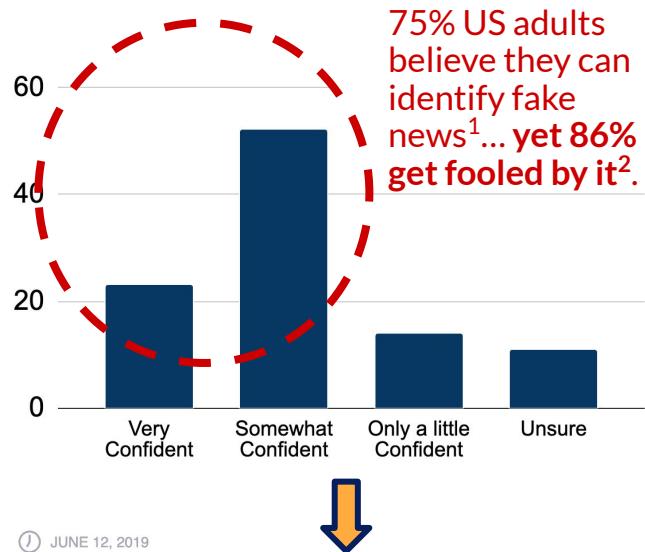


(i) JUNE 12, 2019

86 percent of internet users admit being duped by fake news: survey

¹Statista, 2023 ²Ipsos, 2023 ³Pew Research, 2023 ⁴Ipsos, 2024

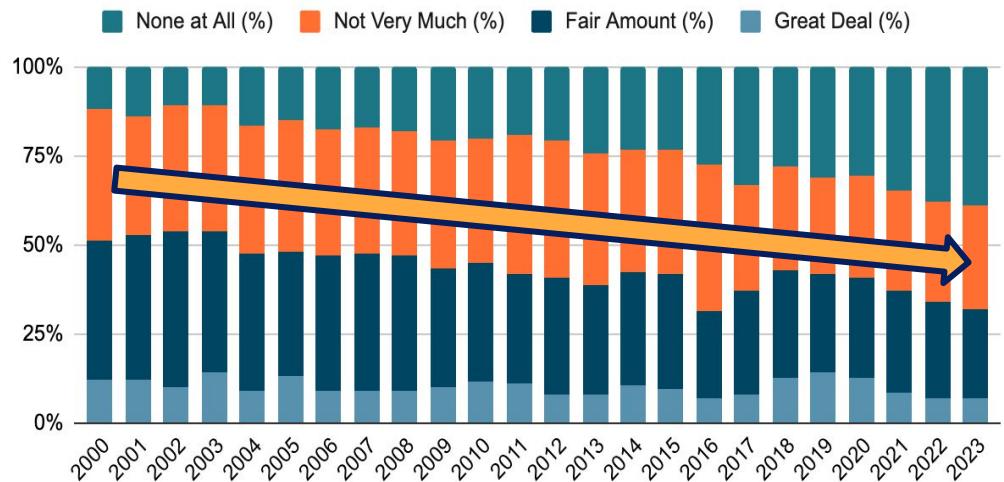
Online Trust Declining: The Misinformation Effect



JUNE 12, 2019

86 percent of internet users admit being duped by fake news: survey

Level of trust in the mass media to report the news accurately among adults in the United States from 2000 to 2023



¹Statista, 2023 ²Ipsos, 2023 ³Pew Research, 2023 ⁴Ipsos, 2024

We build AI tools tackling the Misinformation “Value Chain”

Identify

Threat actors and networks promoting misleading and inaccurate information.

Welcome to Parrot!
A tool for identifying potentially manipulated narratives from state-run media
Click on Open Report to calculate an article's narrative manipulation score. For more information about how it all works, See our [About page](#).

Start here

This narrative concerns: US-Funded Bolos in Ukraine

Russia makes new anti-Ukrainian financed Bolos in U...
Open report

China urges Pentagon to stop supporting rebels in U...
Open report

US-Afghan talks in Ukraine dealt with bat...
Open report

Nazi-collaborator Stepan Bandera
Ukrainian neo-Nazis march to celebrate life of Nazi collaborator Stepan Bandera, a former leader of the Organization of Ukrainian Nationalists

How Western Press Failed to Report on War in Ukraine
Open report

22 million Soviet citizens lost their lives during the Great Patriotic War
Open report

How US state media colluded with US...
Open report

US escalated Ukraine-conflict
Russia criticizes US of escalating tensions by incorrectly claiming Russia is amassing troops at the border of Ukraine and planning an invasion

How Kestrel Steel Year 1 on War in Ukraine
Open report

Russia says military action in Ukraine highly...
Open report

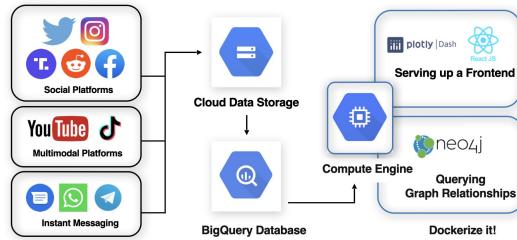
What has been destroyed without a...
Open report

Peace with US only...
Open report

Identify Coordinated Networks

Measure

Influence and effects of misleading information sharing across platforms.



“Social Listening” across Platforms

Mitigate

Harms through interventions to counter the spread of misleading information.



Digital Literacy using GenAI

We Advance Tech and Policy Research

Misinformation Networks

Can we identify misinformation spread by networks of actors?

What is the impact of cross-platform information sharing on social networks?

- ❖ Twitter (preprint, [blog posts](#)),
- ❖ Meta ([report](#)),
- ❖ YouTube (preprint),
- ❖ Telegram ([Stanford T&S '24](#))

User Behavior

Do decentralized platforms increase user agency?

Can a platform's conversations transcend its political affiliations?

- ❖ Truth Social (preprint),
- ❖ Bluesky ([ICWSM '24](#)),
- ❖ Wikipedia ([ICWSM '24](#))

Social Media Policy

How can policymakers regulate social platforms effectively?

Can we use a shared language to discuss different platform interventions?

- ❖ Transparency regulation ([ACM DGO '24](#)),
- ❖ Platform Interventions (preprint)

Safety by Design

How does the manipulation of recommendation algorithms affect online discourse?

Can we limit fake news without censorship and central authority?

- ❖ Social Network Simulation ([ICML '23](#));
- ❖ Designing Marketplaces to Limit Misinformation ([IC2S2 '24](#))

Team



Swapneel Mehta, Ph.D.
Founder and President

Postdoctoral Associate at Boston University and MIT Sloan. Visiting researcher at Oxford, Twitter, Meta, Slack. **Google, MIT, NYU, and Wikimedia** awardee.



Dhara Mungra
Treasurer

Data Scientist at **Bombora** working on Audience Analytics. ML for Finance and NLP Research. M.S. Data Science at **NYU**. Black belt, marathoner.



Harshit Mehta
Secretary

Program Manager at **Maxwell Energy Systems** managing development and delivery of products at scale in the global EV automotive sector.



Raghav Jain
Research Lead

Incoming Ph.D. student at **UCSD**. Tools to fight financial misinformation at Univ. of Manchester, bot detection at **Oxford and Meta**, and LLMs at **IIT and Microsoft Research**.



Board Directors



Dr. Mukund Sudarshan
Co-Founder at Quilt Labs



Dr. Christian Schroeder de Witt
Post Doctoral Researcher,
University of Oxford



Dr. Chirag Raman
Asst. Professor, TU Delft



Sandra Khalil
Partnerships and Trust & Safety
Lead, All Tech is Human



Rebecca Harvey
Online Harms and Trust &
Safety, FCDO (UK Govt.)



Karan Dhabalia
Serial Tech Entrepreneur,
acq. by Stripe



Mansi Panchamia
Developmental Consultant,
World Bank



Stacey Peters
US Local News Leader, Advisor

SimPPL Fellowships to Train Young Innovators

6-9 month programs to train 100+ undergraduate students from the global majority to build and launch responsible computing tools based on our experience founding:

1. **Unicode: Programming Community**
2. **Shalizi Stats Reading Group: Advanced Statistics**
3. **Unicode ML Summer Course: Machine Learning**
4. **NYU AI School: AI/ML education for non-STEM majors**
5. **NYU AI, Misinformation, and Policy Seminar**

Achieved Outcomes:

1. **15+ top tier publications** and invited oral presentations
2. Project partnerships in **6 countries** with **15+ organizations** and agencies
3. **USD 130,000** in competitive global grants; Google and Mozilla



Click to [Review our \(Google\) Fellowships Program](#)



Partners



**United
Nations**



TECHGLOBAL INSTITUTE

Mong Circle



Media and Visibility

- ❖ Invited to **embassies and roundtables** to deliver talks advancing information integrity and artificial intelligence (Finland, UK, Switzerland, UNESCO).
- ❖ Successfully built, deployed, and sold products for data analytics to newsrooms and nonprofits to support underserved communities sustainably, generating **USD 100,000+ annually**.

200,000,000+
Global News Views



\$250,000+
In funding over 3 years



15+
workshop papers



How You Can Support Us

Funders: Financial support (USD 250,000) for 2 years to advance

1. Research / Reports into technology and policy (digital platforms, trust, healthcare).
2. Improving global Responsible AI and Trust and Safety education.

Platforms and Media Houses: Data and Partnerships for up to 3 years to advance

1. Advanced network analysis of cross-platform information and threat actors.
2. Human-centered design of GenAI products and trainings for staff.

Professionals: Pro Bono monthly Mentorship / Feedback to advance

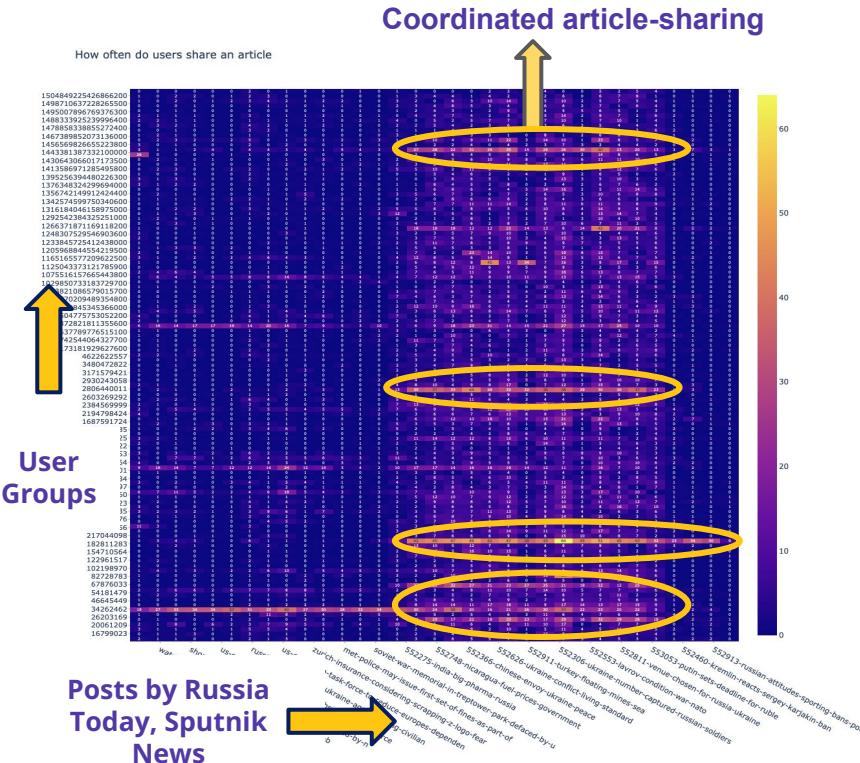
1. Engineering and System Design for AI, NLP, Cloud Software.
2. Sociotechnical Systems, Policy, and AI Governance Research

Sim**PPL**



Research Highlights

Identifying Networks Promoting Misleading Information



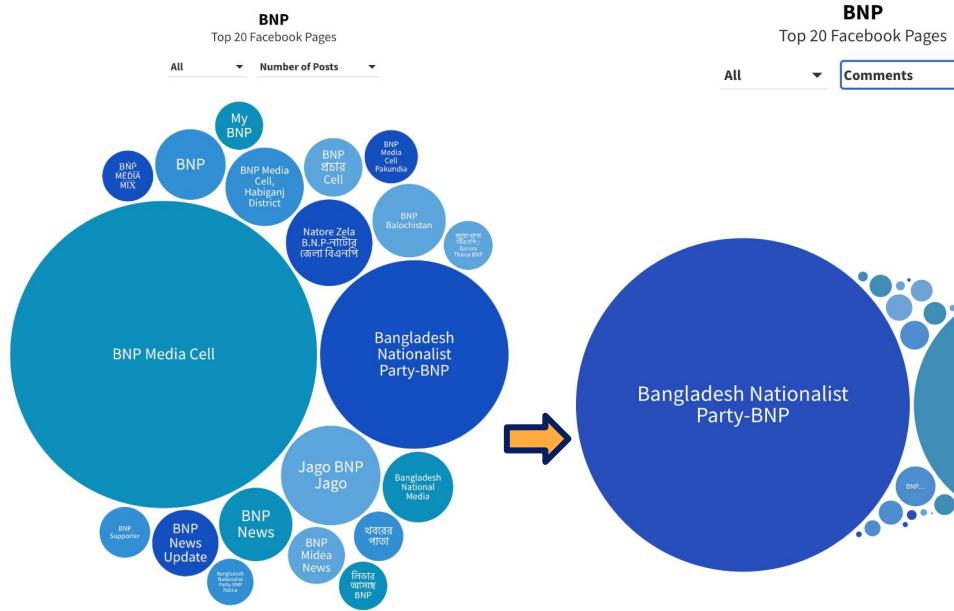
Heatmap represents the account-level frequency of sharing RT and Sputnik articles on Twitter.

70M tweets, 14M accounts over 1 year

Twitter/X's Site Integrity lead reached out to follow-up and analyse the identified accounts ([check out our demo](#), [report](#)).

Won award from Wikimedia to expand project ([feature](#)) into cross-platform coordination detection tools.

Identifying Political Targeting, Gendered Harassment on Meta



Lots of groups and pages posting content (95M+ audience)

But audience attention is focused on a subset of groups making moderation feasible!

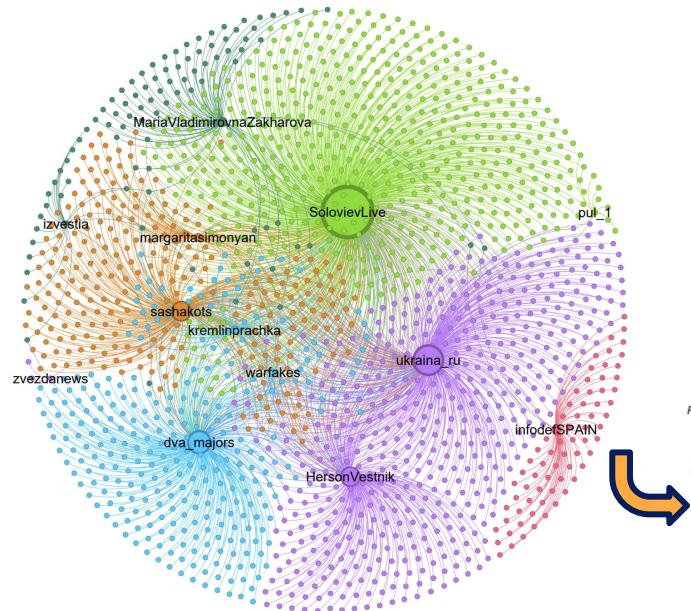
Identifying targeted hate and harassment strategies used to attack political figures in Bangladesh.

600 Pages and Groups, 95M views, 500K posts

Featured in **national news headlines, parliament**, multi-part report into coordinated information operations.

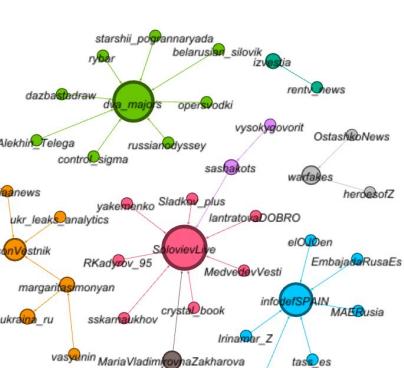
Resulted in **Meta's takedown** of Bangladeshi networks! (**headline, news, Meta's takedown**)

Finding ‘hub’ channels that disseminate Disinformation



4.5k channels spreading pro-Russian disinformation about Ukraine

**Also mobilized to spread
Moscow blast conspiracies**



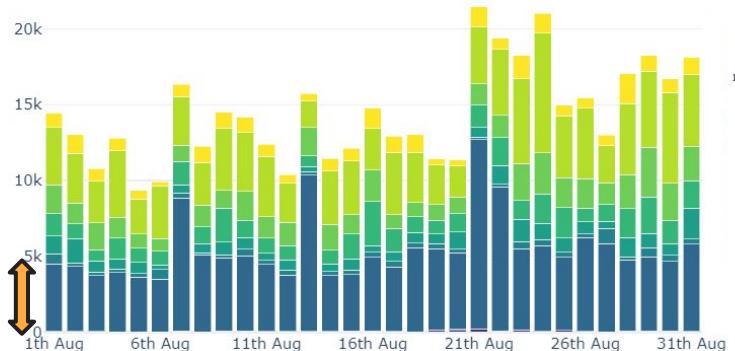
Russian disinformation machinery reuses the same Telegram channels to promote different campaigns.

Investigation in collaboration with
former intelligence agency
personnel accepted at both

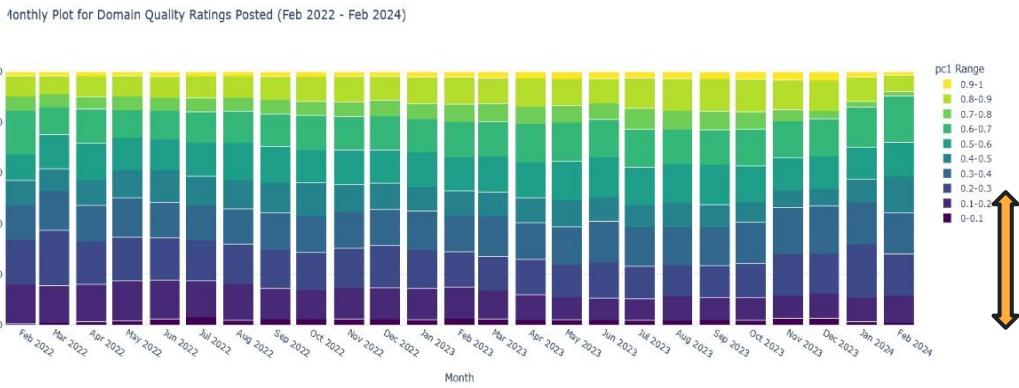
Underground Economy Conference,
Stanford T&S Research Conference

Demo Available!

Longitudinal comparison of Media Diets across Platforms



Much lower volume of low-quality information on Bluesky (Aug 2023)



Much higher volume of low-quality information on Truth Social (2022-24)

Evaluating Influencer strategies on Alt-social Platforms



**Gets shared
a lot**



***Resized for viewing.**

What causes low-quality information to spread on Truth Social?

~2M links shared including 11,000 domain quality ratings over 2 years

Contributed to understanding of how Wikipedia links are shared on Truth Social among pockets of rational discussants ([ICWSM '24](#))

Building: Dashboard to track hourly engagement with elections posts by influencers on the platform.

Sim**PPL**



Product Highlights

Parrot

Use graph based signals to identify coordinated networks on large social networks. Used to identify Russian misinformation around the Ukraine war shared on Twitter at scale across 10M+ accounts. **Won an award from Wikimedia and grant funds from the Times and the Sunday Times.**

Key Features

- ❖ Coordinated link-sharing combined with temporal analysis.
- ❖ State-of-the-art ML/NLP classifiers for “propagandist” content signals.
- ❖ Advancing industry benchmarks on bot detection!

The screenshot shows the Parrot web application's homepage. At the top, there is a navigation bar with links to 'Home', 'Dashboard', and 'About'. Below the navigation bar, a banner reads 'Welcome to Parrot! A tool for identifying potentially manipulated narratives from state-run media'. It also includes a note about calculating narrative manipulation scores and a link to the 'About page'. The main content area is titled 'Start here' and features three columns of news items. Each item has a thumbnail image, a title, a date, and a blue 'Open report' button. The first column is titled 'Biolabs' and includes articles like 'Russia makes new claims on US-financed biolabs in UK...' (07 Mar, 2022) and 'China urges Pentagon to open up about 'biolabs' in U...'. The second column is titled 'Nazi-collaborator Stepan Bandera' and includes articles like 'How Western Press Has Kept Silent For Years on War in Donbas & Neo-Nazism in Ukraine' (26 Feb, 2022) and '27 million Soviet citizens lost their lives fighting...'. The third column is titled 'US escalated Ukraine-conflict' and includes articles like 'Russia says military action in Ukraine highly likely' (02 Dec, 2021) and 'Peace with US only possible after 'new world order'...'. Each news item has a small thumbnail image to its left.

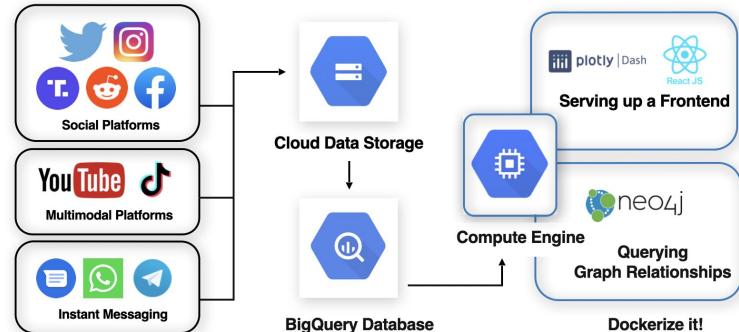
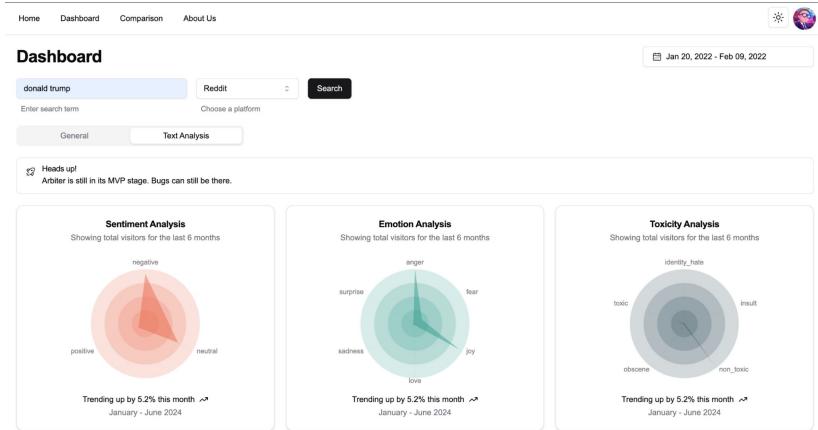
Public Demo Available (site under migration):
parrot.simppl.org

Arbiter

Platform analyzing over 250M+ posts from multiple social networks to measure the networked spread and influence of key narratives, articles, actors, and popular public figures across the information landscape.

Key Features

- ❖ Cross-platform Hashtag & Link analysis to identify trends and pattern of dissemination of information
- ❖ Channel wise statistics to monitor reach and impact
- ❖ Engagement Analytics based on comment analysis for shared links
- ❖ MBFC grading for shared content to understand platform biases and audience preferences



Email team@simppl.org to receive access to the platform demo.

Newsroom Analytics

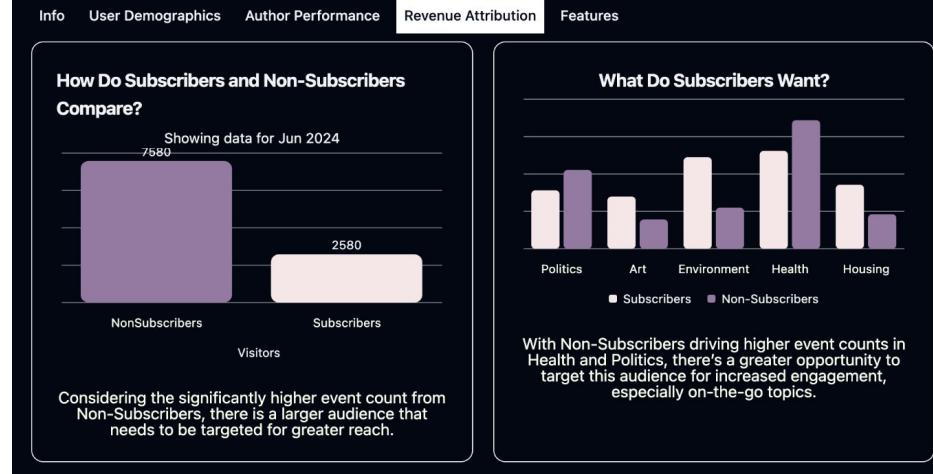
Use GenAI to identify opportunities for content creation and increased audience engagement, elevate 'outperforming' authors across all categories. **Pilot delivered to the New York Public Radio** and expanding to LION Network members.

Key Features

- ❖ Identify audience preferences for news categories using analytics data.
- ❖ Monitor author performance across different categories to increase value to newsroom.
- ❖ Unlock continuous value from decades of user activity data with **conversational intelligence!**

Newsroom Analytics

We help newsrooms make informed decisions, from identifying where to make their next hire to optimizing advertising spending, all aimed at enhancing audience engagement and driving revenue growth.



Our newsroom analytics visual reporting
(now with AI chat available!)

Email team@simppl.org to receive access to report from the 6 month long case study for the WNYC's Gothamist Platform.

Sakhi

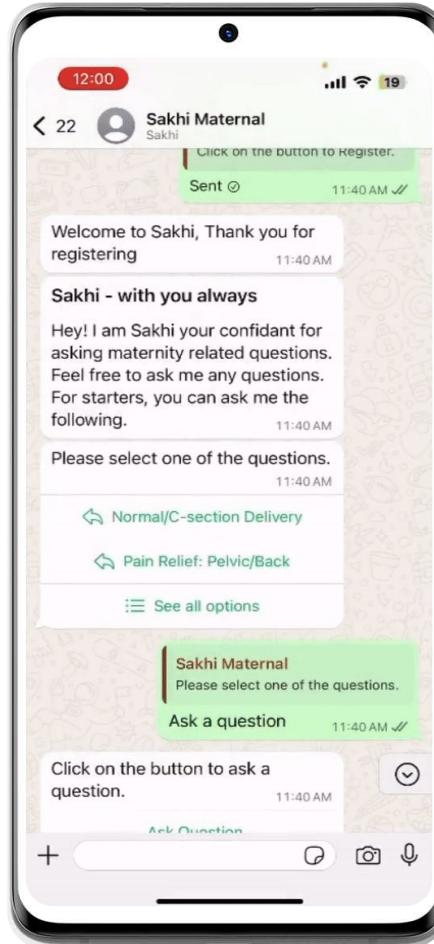
Digital literacy platform designed to deliver accurate and verified maternal and menstrual health information to the mobile phones of users, in multiple languages³.

Pilots ongoing with 300 families in Bangladesh and India, featured in Rest of World.

Key Features

- ❖ Mobile based multimodal messaging.
- ❖ Monitor last-mile care delivery programs
- ❖ Gamified rewards for engagement.
- ❖ Mitigate the effects of knowledge gaps and misinformation.

³Sakhi. For chatbot, scan the QR code and email swapneel@mit.edu to receive a password for access.





Talks and Presentations Featured Globally



HOW DID TWITTER'S INTERVENTIONS AFFECT TRUMP'S CAUSAL EFFECTS

- Estimating the causal effect of Twitter's interventions on Trump's Tweets
- Two types of interventions: **warning labels** and **removal**
- Looking at multiple platforms reveal surprising conclusions about the interventional effects



UNESCO World Press Freedom Day
Deutsche Welle
Jagran New Media
DW Disinformation Learning Guide
Embassy of Finland
Humane Intelligence
All Tech is Human
Rest of World
Swiss Embassy
MIT Media Lab
Stanford Internet Observatory
Computation + Journalism
Columbia University
NYC Media Lab
New York University
McKinsey, Twitter, Graphika, others