

The strategy behind innovative data ad products at the New York Times

Presented by Nick Eckhart

October 19, 2021

1

I work at a small, medium, medium-large, and large business.

4

I live in a condo, apartment, and single-family home that I either own or rent that's valued somewhere between \$0-\$750K+.

2

My functional area is sales, C-Suite, consultant, engineering, finance, legal, marketing, medical/health, and operations.

5

I'm a parent.

3

My credit level is excellent, good, and poor.

6

I purchase a lot of hot sauce.

1 Challenges of Third-Party Data

2 The role of privacy

3 The opportunity

4 Contextual & Insights Products

5 1st Party Behavioral Targets

The New York Times stopped allowing the use of third-party data at the end of Q1 2021.

Questionable Accuracy

Third-party data can approximate an audience, but many times the segments are contradictory or wrong.

Lack of Transparency

Hard to build smarter and more ethical audiences if it's unclear how the data is collected that makes up the segment.

“This category contains users who work as nurses.”

“This segment targets Nurse by the specific qualified job title.”

“Information is compiled from legally approved state professional license files and membership confirming that the occupation title of the 1st individual as Nurse”

- 1 Challenges of Third-Party Data
- 2 The role of privacy**
- 3 The opportunity
- 4 Contextual & Insights Products
- 5 1st Party Behavioral Targets

Cross-Context Identifiers Are Disappearing

What Are Cross-Context Identifiers?

They are ways of recognising users as being the same person not just inside of a single context (site, app) but between contexts.

Which Cross-Context Identifiers Exist?

There are many, but the primary ones are mobile advertising identifiers, known as “MAIDs” (sic), and third-party cookies.

What Are Cross-Context Identifiers Used For?

For many use cases in which it is useful to keep track of a person’s behavior as they move between places:

- Behavioral targeting (user reads content in places A, B, and C, targeting them for it in D)
- Frequency capping (user has seen that ad elsewhere five times already)
- Conversion & Attribution (user converted after seeing the ad on site A and clicking it on site B)
- And More: the industry has used this every which way



All Major Identifiers Are Likely to be Affected

Apple IDFA (Done in 1H21)

IDFA is Apple's Mobile Ad ID. It is an ID tied to a mobile phone and is used to track user activity across different apps. It was created for advertising purposes.

In 1H21, Apple iOS will no longer allow marketers to access mobile device identifiers without explicit user consent.

With iOS accounting for 46% of all mobile phone usage in the US, we expect the addressability or measurability of app media to be drastically reduced.

Browser Cookies (1H23)

A cookie is a piece of data that is tied to a user's browser. Third-party cookies are used as identifiers to track people across different websites.

Originally 1H22 and now 1H23, Google Chrome, which accounts for two-thirds of the browser market will no longer support third-party cookies. Safari and Firefox, which account for the remainder, already block third-party cookies.

Google will not use any identity solutions, such as those that depend on hashed emails, in its own buying/selling platforms, but it will support a mechanism for publishers to pass identifiers to third-party buyers.

Google AdID (End of '21)

AdID is Google's Mobile Ad ID. It is an ID tied to a mobile phone and is used to track user activity across different apps. It was copied from Apple for advertising purposes.

Google already allows Android users to opt-out of personalized ads, but starting later this year the AdID for opted out users will no longer be available. We expect more restrictions in the future.

AdTech Is Undergoing Disorganised Reform

Some Are Reinventing Cookies

Several “unified” ID schemes are competing to rebuild cookies from hashed emails. This does not align with user expectations, Apple already has counter-measures, and others will follow.

Cohorts Targeting

One of Google’s proposals, FLoC, has browsers cooperate to analyse user behaviour across sites and place them into “cohorts” (ie. segments) pitched as privacy-friendly. This uses publisher data across contexts and puts people in groups using opaque machine learning.

Many “Bird” Proposals

Numerous “bird” proposals (TURTLEDOVE, FLEDGE, PARAKEET...) from varied sources seek to introduce features such as retargeting, anonymisation for contextual ads, or frequency capping.

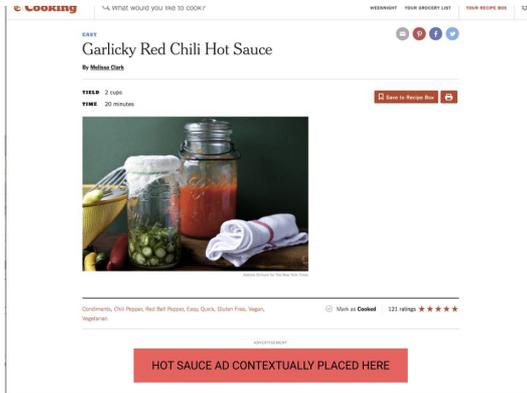


- 1 Challenges of Third-Party Data
- 2 The role of privacy
- 3 The opportunity**
- 4 Contextual & Insights Products
- 5 1st Party Behavioral Targets

The New York Times mission of seeking the truth and help people understand the world drives our ability to create innovative, differentiated products

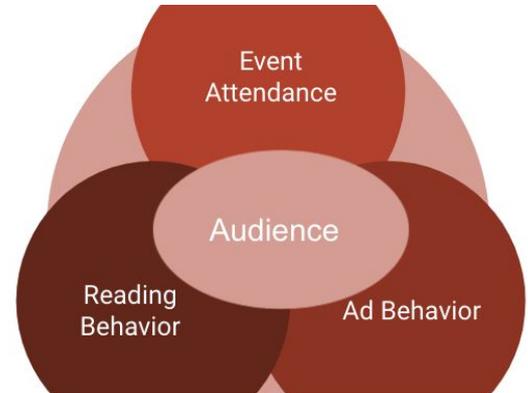
- 170th anniversary celebrated this year 🙌
- 54,000 stories published in 2020 📖
- 100 million registered users 🙌
- 8 million paid subscribers 👁️
- NYT Readers have a greater expectation of privacy
- Hire talent in essential areas: Data Governance, Data Science, Product Marketing, and Data Engineering 🤝

The Opportunity



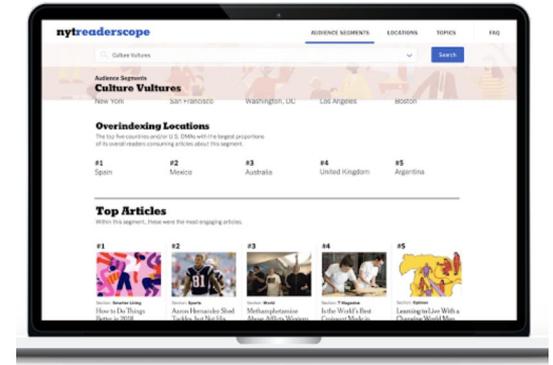
Contextual Products

Targeting users based on the page context



Behavioral Products

Targeting audience groups based on behaviors they've exhibited online or offline

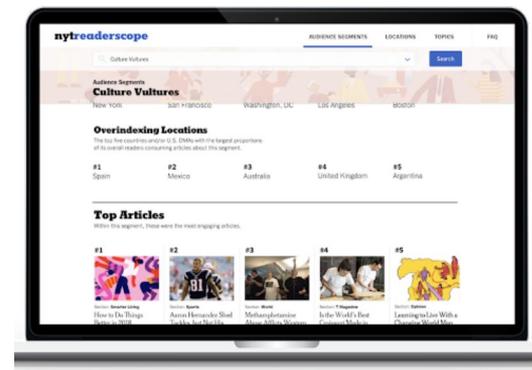
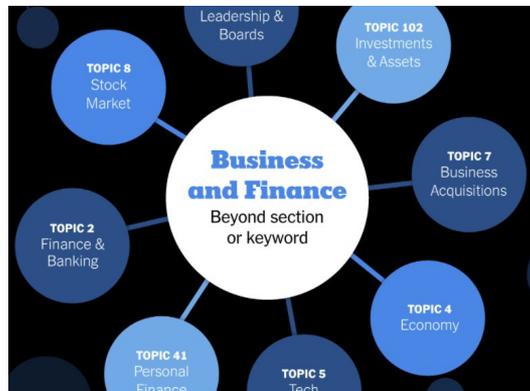


Insights Products

Marrying what we know about the user with what we know about the content to help clients strategize

- 1 Challenges of Third-Party Data
- 2 The role of privacy
- 3 The opportunity
- 4 Contextual & Insights Products**
- 5 1st Party Behavioral Targets

Contextual & Insights Tools



Perspective & Motivation Targeting, which are predictive models on how the article might make a user feel

Topic Targeting: Most articles fall into multiple sections. This product uses a Latent Dirichlet Allocation (LDA) topic model to allow advertisers to reach topics more broadly than section.

Readerscope, which is an internal insights tool that marries the deep knowledge of our users and content to help inform campaign strategy

Products in Action

Perspective Targets

- Inspired
- Adventurous

Motivation Targets

- Go on an outing
- Make a Healthy Change

Topic Targets

- Nature
- Wellness

Readerscope

- Most read by “Millennials who like to Leisure Travel” and “DIY’ers”

The New York Times



- 1 Challenges of Third-Party Data
- 2 The role of privacy
- 3 The opportunity
- 4 Contextual & Insights Products
- 5 1st Party Behavioral Targets**

Demographic Segments

1: Proof of Concept

- a) Would users would be willing to give us some information knowing it will be used for Advertising?
- b) Are we are able to use that information to create reliable models?

2: Collect Data

Prioritize the most important segments and determine how to collect enough responses

- a) Quarterly survey to regi/subscribed users
- b) Internal promotion spots

3: Explore Models

We match the over 75K people who have responded to our survey with over 900 onsite behaviors e.g:

- a) Devices used
- b) Articles read from a particular desk
- c) Where you are in the subscription funnel

Demographic Segments

4: Validate Models

It was important to us to ensure the segments performed better for clients.

- a) F1 score
- b) A/B Tests
- c) External Audit

5: Delivery Pipeline

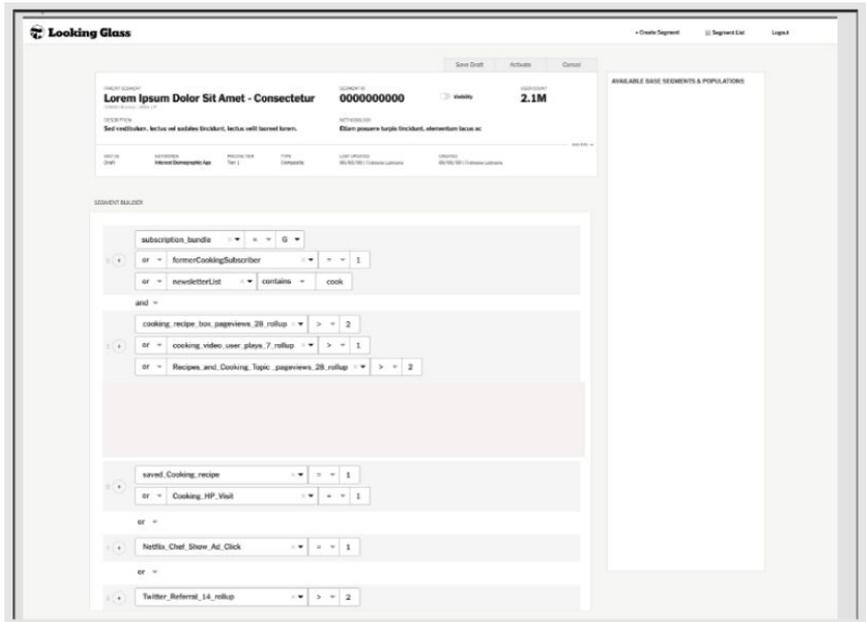
Once we knew the models were performant, we had to build a complex workflow including:

- UI to build segments
- Daily scoring job that put users into segments based on the rules created in the UI
- Pass those scores through internal systems to ultimately end up in ad call for targeting

Six: Iterate

- Monitor & Improve CTRs
- Build new segments per market demand
- Using other data such as clicks to ads to create models

Behavioral Targeting Interest Segments



Looking Glass

A UI to allow us to build segments

The New York Times

Custom Topics Parameter Search Engine

An app that allows the user to find parameters (search queries, exclusionary words, and similarity score cutoffs) for use in custom topics.

Search:

How many articles do you want back?

n =

Top Articles

Exclusionary Words:

Exclusionary word group 1:

Exclusionary word importance (0-1)

Exclusionary word group 2:

Exclusionary word importance (0-1)

Custom Topic Builder

A UI to hone in on particular topics

How are we doing?

The Good

-We can tell advertisers exactly how our segments are built

-Significant savings in data costs by no longer using third-party data

-74% increase in click-through rate for behavioral targeting

Challenges we've faced

-Publishers that still use third-party data can offer more niche segments (specific titles, company size, etc.)

-Maintenance: Getting a continual flow of new survey responses, retraining the models each quarter, and ensuring all of the systems are working properly is a large body of work