

Helping small and local newsrooms harness their superpower

Written By Ariel Zirulnick and Becca Aaronson

For the news industry to lead in the AI era instead of chasing it, publishers need a first-party data infrastructure that enables even the smallest newsrooms to achieve better product-market fit and enduring revenue streams while utilizing AI.

Although AI poses a threat to existing business models, when paired ethically and strategically with first-party data, AI can also help publishers develop more resonant journalism and editorial products, strengthen relationships with targeted distribution strategies, and build more resilient revenue streams.

For example, insights from first-party data would enable them to improve:

- **Editorial strategy:** Spot underserved topics or communities and apply those insights to editorial strategy to develop more resonant stories and products, such as organizing an event in a neighborhood that is engaging with your journalism at unusually high levels
- **Distribution:** Increase loyalty by personalizing newsletters, homepages, offers, and other audience-facing products, such as sending your loyal education coverage readers a direct invitation to a school board candidate forum you're hosting
- **Revenue:** Strengthen ad sales by giving local businesses precise, privacy-safe audience matches, such as showing a local coffee shop that 20% of your "morning newsletter regulars" live within a mile of their location

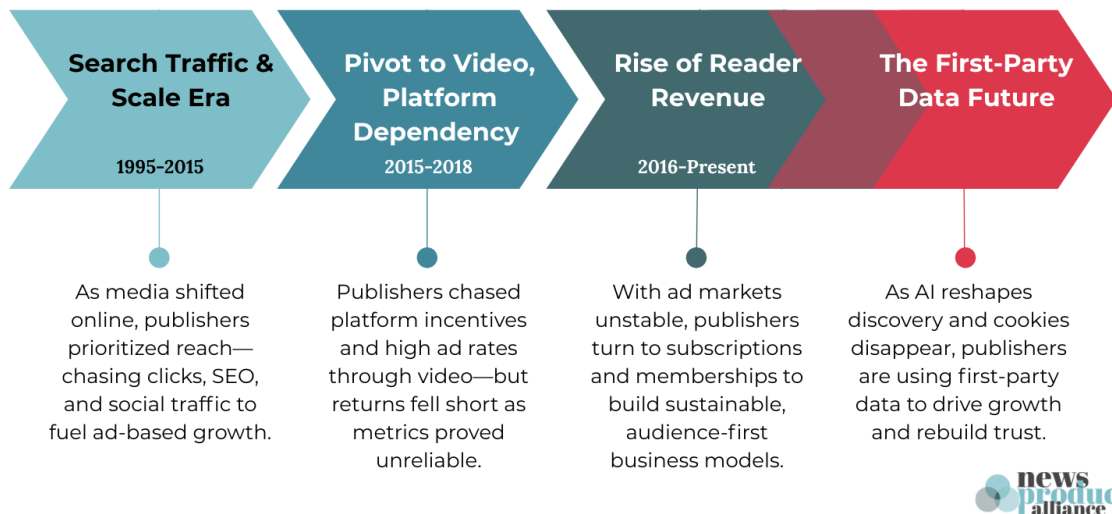
These wins are only possible with strategic collection, segmentation, and application of first-party data. Most publishers are aware of this potential, but need help getting there. According to Omeda's [2025 State of the Audience report](#), 85% of publishers agree that audience data is a competitive advantage — yet only 45% routinely update strategy based on it, and only 36% regularly use it to personalize or innovate. As Omeda put it: “The era of passive data collection is over — now it’s about making it work.”

We’re far from the first to flag the power of first-party data. Many B2B publications have operated on a first-party data model from inception because their model relies on knowing exactly who they’re reaching. Jacob Donnelly, author of “[A Media Operator](#),” and Brian Morrissey, author of “[The Rebooting](#),” have been tracking this shift for years, highlighting the success of early adopters like Bloomberg and Dotdash Meredith.

In [a recent Rebooting webinar on how and why Forbes unified its data](#), Morrissey provided a clear overview of the five macro trends driving the need for publishers to understand their audience and “activate that against their business model”:

- The decline of the pageview economy, which shifts publishers’ focus from volume to value
- The [collapse of search](#) and increasingly fragmented distribution, requiring publishers to earn more from smaller audiences.
- The pivot to build and own direct audience relationships in response to the discovery and distribution collapse
- The “great refactoring” of business models—building the next one while still running the old
- The need to act like a brand, selling access to audiences, not just content.

The Evolution of Digital Media Revenue



Many newsrooms have already taken steps in response to these trends, focusing on email collection, site registration, and app strategies to rebuild direct relationships that suffered from years of search and social reliance. (You'll often hear this described as an effort to convert anonymous users into known users, or A2K.)

That's first-party data: the information you collect directly from your audience — who they are, what they care about, and how they interact with you. For most newsrooms, the anchor for first-party data is an email address. That's the key that connects a person to their engagement history, location, and interests. The next step is connecting that data to product development —utilizing that useful, actionable information about each person to build news products that deliver more value.

For small and local publishers, it's a key to sustainability. Because without a strategy for collecting and utilizing first-party data, news publishers risk being outpaced by competitors (both in and beyond news) who can be more responsive to their audience's needs, being invisible to advertisers who increasingly demand proof they can reach advertisers' targets, and being irrelevant to funders who expect impact evidence. In short: you don't get to opt out.

Of course, first-party data can't solve all audience and revenue challenges. It doesn't solve the discovery challenges created by the decline of search and social, for example. But first-party data and improved data operations can help publishers understand what content, formats, features, and platforms attract, engage, convert, and retain audiences so that they can make their offerings more compelling.

When the AI Co-Lab [began its work to help small and local newsrooms better use their first-party data](#) to understand, segment, and reach their audiences, we expected to encounter technical debt and data management challenges. But the [gaps in expertise and technical capabilities](#) were more staggering than we had imagined. It's clear that as national publishers [invest heavily in shoring up their first-party data infrastructure](#), small and local publishers are being left behind.

But small and local publishers have a superpower — trust and proximity to their communities — that can secure their sustainability if paired with a smart first-party data strategy.

The AI Co-Lab plans to address the expertise and technical gap in three ways over the coming months:

- Practical guides and blog posts on topics such as collecting, structuring, and applying first-party data; ethical considerations; and rethinking sales strategies
- A common data schema for consistent and shareable audience insights
- Open-source tools to help publishers centralize and use their data effectively

What becomes possible when small and local newsrooms align collection, segmentation, and application of first-party data and get an assist from AI to do that at scale? That's what we want to find out – and enable.

A Quick Guide to Audience Data Types

- **Third-party data** is collected by outside companies and shared with others, such as data aggregators, data brokers, or advertising platforms. It's broad, often aggregated, anonymous and outdated, and increasingly restricted by privacy regulations.
- **Second-party data** is collected directly from users by another organization and shared with you. It's someone else's first-party data that's shared in a partnership.
- **First-party data** is collected directly from users on owned channels, such as websites, apps, newsletters or events. It can include both known and anonymous users, depending on the gating strategy of the publisher.
- **Zero-party data** is provided intentionally by a user to a company, like survey responses or preference settings. It's often considered a subset of first-party data, and also referred to as declarative data.

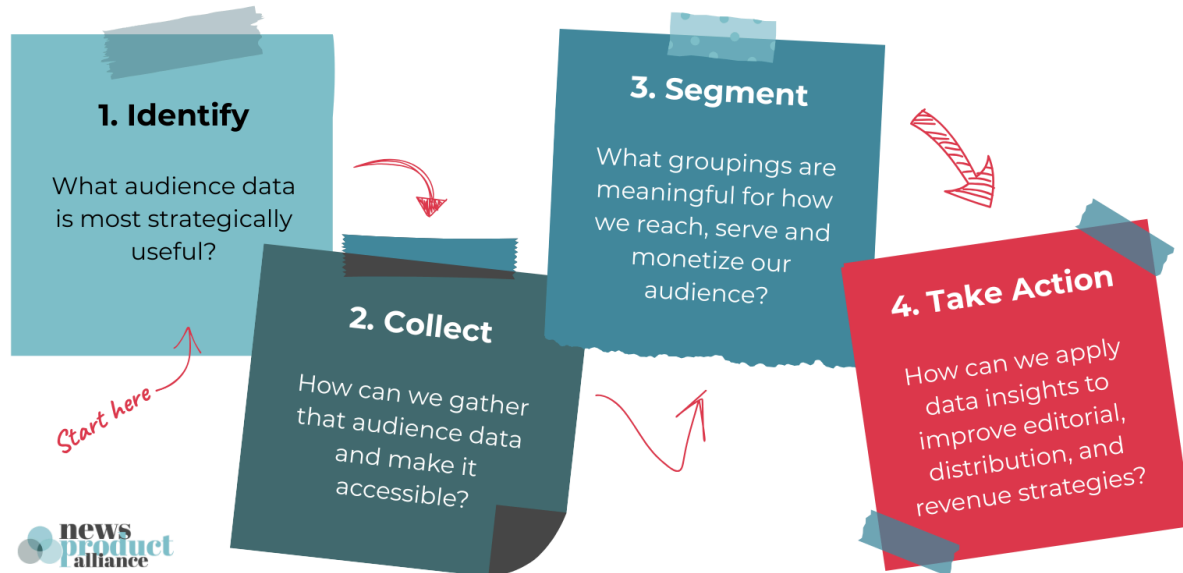
Credit: Kevin Charman-Anderson

What does a news organization that effectively utilizes first-party data look like?

Most conversations about first-party data focus on the application after a story or product is created, using it to fine-tune marketing, distribution, ad targeting, and personalization. That's valuable, but it leaves an enormous opportunity untapped.

At NPA, we believe first-party data is just as critical before a story or product is developed. It should be treated as a steady stream of audience insights that help align your journalism with audience needs, sharpen product-market fit, *and* lay the groundwork for smarter, more effective distribution and monetization.

How to use first-party data in a news organization



If your stories and products aren't informed by audience insights, no amount of growth hacking will make them truly stick.

This is no easy feat. Newsrooms struggle to collect and manage the kind of rich, declarative data that makes this possible, let alone connect it on an ongoing basis to their editorial strategy. Improving this process is one area that the AI Co-Lab has zeroed in on for support. (More on that toward the end of this piece.)

But there *are* newsrooms who have found ways to move forward despite that challenge, including:

- [Education Week](#), which covers K-12 education news and information. They ask readers to share their connection to education (six categories, from teachers to policymakers) so it can target content and offers accordingly.
- [Technical.ly](#), which covers entrepreneurship and jobs of the future in emerging U.S. tech hubs. They identify whether someone is an entrepreneur or technologist, and what stage of

career they're in, which they use to tailor coverage, editorial products, events, and advertiser offerings.

- [El Tímpano](#), which serves Spanish and Mam-speaking immigrants in the Bay Area via SMS. They collect data such as languages spoken, zip code, and whether subscribers have young children in the home to ensure their messages are both relevant and actionable.
- [Village Media](#), which owns a network of local publications in small and mid-sized cities across Canada. Custom-built tools such as Village IQ, their polling software, and Spaces, their local social media platform, give them city-level insights into people's interests. Those insights inform editorial strategy and also allow Village to offer more precise advertising opportunities to local businesses for which they anticipate being able to charge higher rates.

None of these newsrooms collect this data from *everyone* in their audience, but they collected it from enough of their users to form a meaningful sample that can be used to inform editorial, product, and revenue strategies.

If you're in a small or local newsroom experimenting with first-party data, we want to hear what you're trying and what you're learning along the way. Reach out to Ariel at ariel@arielzirulnick.com.

The process for successfully leveraging first-party data looks something like this:

1: Identify what data is most strategically useful

The quality of your audience insights depends on the quality of the data you collect.

Your strategy for collecting first-party data should be guided by your organization's mission, scope, and goals, as it is in the newsrooms we mention above. If you want to shape stronger products, deepen

engagement, and open new revenue opportunities, you need to gather information that directly supports those goals — not just whatever happens to be easy to capture or is in the vendor’s default settings.

Start by asking:

- What do we need to know to serve people better?
- What would help us reach them more effectively?
- What would make our offers more compelling to funders, advertisers, or subscribers/members/donors?

Even general-audience publications benefit from defining which audience identities are most relevant for product development and distribution — because every “general” audience is actually made up of many smaller audiences with different characteristics and needs.

2: Collect that strategically useful data

You can gather data at any point where you exchange a service for an email address — newsletter signups, event RSVPs, and account registrations, for example.

Start small by adding one or two questions to existing forms, making sure not to introduce so much friction that you cut into completed registrations or transactions. Consider using surveys to periodically capture richer detail from your most engaged audiences.

When your organization has a shared understanding of which audience data is most valuable, you’re more likely to spot collection opportunities across departments. You may also realize you already hold a wealth of untapped information. Simply naming the data that is strategically useful can reveal new value in what you’ve previously gathered.

Tip: *Whatever you collect, it needs to flow into a shared database accessible across teams – not stuck in separate spreadsheets, Mailchimp lists, or*

survey tools. Data stuck in siloes can't reveal the full picture of your audience. We have a guide to centralizing your data coming soon.

3: Create audience segments

Segmentation is the process of grouping your users based on meaningful characteristics so that you can take action. [As Emily Goligoski and Emily Roseman wrote in the Membership Puzzle Project's Membership Guide:](#)

Here's why segmentation is so important: your audience is not a monolith. It is composed of different sub-communities or segments, and until you discover those segments, you don't really know who your audience "is" and you can't take any action. When you can accurately segment your audience, **it means you know something vital about them.**

Goligoski notes that these are the four categories of data that are used the most often to segment audiences:

- Behaviors (what users do)
- Attitudes (what users think)
- Demographics and identities (who users are)
- Geographies (where users are)

There are countless ways to segment an audience, and how you do it depends on what you're trying to accomplish. If your goal is to target loyal readers in a membership campaign, you might segment based on recency, frequency, and volume of interaction with your organization. Audience-centric newsrooms typically go further, layering on insights into interests and motivations – for example, if you want to promote an event, you might segment based on topical interests related to the event.

Why is first-party data valuable for news organizations?



Editorial Strategy: Identify audience needs and interests to develop or improve editorial products — *e.g. organizing an event in a neighborhood that's engaging with your journalism at unusually high levels*



Distribution: Deliver the right products to the right people through personalization or targeted distribution — *e.g. sending loyal readers of your education coverage a direct invitation to a school board candidate forum you're hosting.*



Revenue: Prove your reach to advertisers and sponsors by building clear audience segments — *e.g. showing a local coffee shop that 20% of your "morning newsletter regulars" live within a mile of their location.*



4. Apply segments

Collecting and segmenting data is only half the job. You only get a return on that investment when you *take action* with those segments – when you move from knowing something actionable about your audience to doing something with that knowledge. That could mean:

- Developing a new editorial product based on the needs of a specific audience segment – and getting that product sponsored pre-launch because of its grounding in a data-backed articulation of that product's value proposition
- Adjusting how you package and promote existing content to resonate with different groups (while newsrooms are often drawn to making *new* products, NPA sees a big opportunity in using first-party data to redistribute evergreen or underseen stories to the people most likely to find them valuable)
- Pitching advertisers and sponsors based on your ability to guarantee delivery to audiences that align with their goals (this is the foundation of El Tímpano's [civic partnerships strategy](#), now their second largest source of revenue)

Small and local publishers are behind the curve when it comes to leveraging first-party data to secure advertisers and sponsors. But national publishers have already demonstrated the ROI. When Bloomberg turned off third-party, programmatic ads back in 2022 and started targeting ads based on its own first-party data instead, [CPMs shot up by 20%](#).

Small and local newsrooms might not have the technical capabilities to do on-site activation like Bloomberg, but almost any newsroom can experiment with email. Using first-party data to target sponsored emails has two benefits – you can experiment with charging higher rates based on the tailored audience and ensure you don't exhaust your whole audience with messages that aren't relevant to them.

When done well, applying audience insights creates a virtuous cycle: your journalism becomes more relevant, which deepens loyalty, which increases the amount and quality of first-party data shared – making your next products, distribution efforts, and monetization opportunities even stronger.

Where AI fits into this

Everything above can be done manually – but for most small teams, that's slow, inconsistent, and hard to repeat. AI can help you do the same work faster, more often, and with fewer resources, turning first-party data from “something we look at once in a while” into a tool for regular decision-making.

Here are three ways it can make a difference:

Structuring unstructured data: The richest audience signals often live in harder-to-wrangle places – survey responses, reader emails, comment threads, audience questions, and direct emails. Without a comprehensive system for organizing and analyzing that feedback, it remains scattered and underused. AI can rapidly turn that unstructured

input into organized, actionable data, and surface themes you can use to shape editorial decisions, product development, and audience strategies.

Dynamic segmentation: Static “everyone gets everything” lists limit your ability to be relevant. Dynamic segments let you reach exactly the right people for a given story or product — even on deadline. AI can instantly create or refresh those segments based on your parameters.

Personalization at scale: Personalized products build loyalty, but manual personalization is rarely sustainable. AI can adapt content for different audience segments without requiring separate production workflows.

The bottom line: First-party data is the foundation. AI is the accelerator. When small and local newsrooms combine them, they gain the ability to deliver the right journalism to the right people — and build sustainable revenue around it.

Kevin Charman-Anderson contributed to this post.