A Guide to Working With People's Data *Responsibly*

For Strategists & Storytellers to Minimize Risks and Maximize Results from Data-Driven Programs

thicket

How to Cite this Guide

thicket

thicketlabs.com

Foreword

The recent revelations about Cambridge Analytica's exploitation of Facebook data contain serious implications for strategists and storytellers, and any organization exposed through a similar situation may face serious consequences. The richness of the data we use in modern campaigning, and the ease with which we can connect it back to individuals, demands greater awareness and respect for privacy. The process of campaigning has always been about using data to craft sophisticated messaging that persuades people to make a choice regardless of their original position. The current controversy threatens the very notion that such messaging is acceptable behavior.

As a strategist and storyteller who worked on my first digital campaign in 2003, I've had a front row seat for the evolution of the digital data ecosystem that propels our knowledge economy. I've also been trusted with literally millions of people's data at different points in my career. Sometimes in order to get better at positioning or marketing, sometimes in the service of research or the distribution of important information. My story is not unique.

When it comes to working with data, many of us are currently working without a net. Even with the best of intentions, many assumptions that drive business and nonprofit operations are in direct conflict with sound data practices, heighten risk factors for exposure through data, and degrade the quality and applicability of data and insights gathered.

I've put this guide together to help those working with people's data to frame important questions that don't usually get asked in the course of business. It's important to note that this is not a guide for data practitioners with formal training or for those working in specialized fields such as healthcare where data collection is governed by specific policies and regulations. The intent is to offer a practical framework for strategists and storytellers to evaluate their approach to data collection, analysis, and application to minimize risks and maximize quality outcomes from their data-driven programs.

Deepthi Welaratna
April 2018

thicket

thicketlabs.com

thicket

Working With People's Data Responsibly

In This Guide

Data Program *Planning Process*

Step 1: Why Am I Collecting Data?

Data collection should have a purpose and a defined timeframe. Are you doing customer research, testing a product idea, developing a new market? **Identify and write out your goal, audience, deadline, and intended outcomes. This scope will help you limit your data needs, save time and money, and reduce risks associated with sensitive data collection.**

Step 2: Is My Audience on My Team?

Who are you planning to collect data from? A strong data team will have audience representation on it. **Audience representation will help you avoid pitfalls that can degrade the quality of your data or resulting insights.** Your audience is your expert, and you'll need their buy-in. Are there people on your team who reflect audience demographics? If not, invite your intended audience to weigh in. **Include audience input in the design phase of your plan to help you understand what data to collect, how to ask for it with sensitivity, how to reach your intended audience, and how to incentivize participation.** If you do ask audience members to consult on the design of your plan, pay them for their time and expertise.

Step 3: What Data Am I Collecting?

Does meeting your goal mean you must collect people's personally identifying information? Do you need to collect data that puts anyone in danger? Will any of your questions cause trauma or discomfort in your audience? **Make sure you ask yourself: What is the bare minimum of information that I could collect that will achieve my intended objectives? Reducing the amount of data you collect reduces risks associated with the collection of it.** For example, if you don't need to know a survey taker's name, don't ask for it. Challenge your assumptions about how much data is necessary to achieve your goals.

Data Program
Planning Process

Step 4: How Am I Collecting and Storing Data?

Will personally identifying information be kept with individual answers? Is the data being collected passively, for example by monitoring website or mobile activity, or actively through interviews or surveys? **With more passive methods, there is ongoing risk of privacy and trust breaches. Active methods start with gaining a person's express permission.** How and where will this information be stored, and who will have access? Does the data belong to your organization exclusively or is it also shared with a platform being used to collect it? How long will you keep the data and how will you discard it? **Responsible data programs should hold personally identifying data for the briefest time possible.**

Step 5: How Am I Providing Value to My Audience?

This question has to be answered at several different levels throughout the data plan. First, what is the benefit to an audience member who gives their input? **Even if you've decided to collect anonymized data, you still have a chance to build a relationship with your audience that deepens engagement and builds trust. The data you collect may have educational value to your audience** such as giving them insights that can help them learn about themselves or make better decisions.

Step 6: How Am I Analyzing the Data?

How complex is the data, and how large is the dataset? Are you the data analyst? Do you have or need an expert data analyst? Is a spreadsheet analysis enough, or do you need the advanced analytics available in a tool like R? **The analysis is another key step where audience involvement will improve your results.** If you don't have a team member who represents your audience, hire one.

Data Program
Planning Process

Step 7: How Am I Reporting the Data?

How will the data be used for reporting? Are you reporting anonymized aggregate numbers, or will you include individual datasets? Is there any potential for re-identification? Who will have access to the reports? How will the results be used by its readers, and is there potential for exploitation of audience members based on the results? **Think about reporting only the essentials to minimize risks.**

Step 8: How Will This Data Project Impact My Audience?

Assess how you plan to use people's data in the bigger picture: Are you using it for their benefit, or are you using it against them for manipulation?

Persuasion

Will the people who gave their data be more vulnerable to outside influence after participating? Will they be more easily persuaded to make choices that run counter to their own interests?

Behavior Change

Will you be targeting the people who gave their data with a behavior or opinion change program? If the goal is to change a preference or behavior, what are the other implications of that change, now and in the future? W**ill they be more susceptible to starting, maintaining, or accelerating an addiction or unhealthy behavior?**

Empowerment

Will the people who gave their data be in an empowered position after participating? **Will they have learned something about themselves such as knowing more about their own preferences, learning new skills, or receiving insights that help them make better choices for themselves going forward?**

Informed Consent

Some programs are based around collecting publicly available data about people. In fact, companies have created entire viable businesses by aggregating datasets about people that are valuable to a specific industry and monetizing access to that data. Unfortunately, many people don't know that their data is publicly available, or to what extent it's available, leading to a default state of compromised privacy for many.

Regardless of whether you're collecting people's data directly or indirectly, gaining informed consent is a critical and non-negotiable component of working with people's data responsibly. It's important to note that checking the box on a lengthy and inaccessible list of terms and conditions is increasingly not considered to be informed consent.

At a minimum, informed consent means that the person whose data is being collected and applied learns about and understands the purpose, benefits, and potential risks of participating in the data-driven program, including who will have access to their data now and in the future, and how their data will be applied, reported, retained, and/or discarded.

Given the risks associated with keeping people's personal data and the high barrier to gaining informed consent, it should be clear why researchers prefer to design data collection programs to collect as little personal data as possible. But what about for companies that rely on aggregating publicly available data in their daily operations?

Many companies find it tempting to ask for forgiveness, rather than ask for permission, by providing opt-out procedures rather than opt-in. For companies who do this, consult your target audience about how it benefits them to participate in your program and how it affects their safety and security. Evaluate whether your program is exposing you to unwarranted risks.

You may find that it's necessary to scale back the display of data for people who have not yet given consent, or that shifting to an opt-in model will be necessary. You may even find that your program cannot co-exist with responsible data practices. As we move towards a more heavily regulated data environment, it is wise to take action before you are forced to it at a later date and after more time and resource investment takes place.

Audience Inclusion & *Checklist*

Data Collection

Are the people being asked to provide their data weighing in on what data is meaningful to collect?

Informed Consent

Are the people being asked to provide their data fully aware of how their data will be collected, stored, used, and shared -- now and in the future?

Analysis

Are the people being asked to provide their data weighing in on what the data means?

Value

Will the people who gave their data get to learn from it or use it to their own advantage through a report back, a recommendation, or other benefit?

Future Use

If you keep the data and continue to derive value from it, will the people who gave their data also continue to derive a benefit from it?

Impact

Will the people who gave their data be in an empowered position after participating? Will they be targeted with a behavior or opinion change program? Will they be more vulnerable to outside influence?

About the Author

Deepthi Welaratna is the founder and CEO of Thicket Labs, an audience research and innovation company helping clients engage audiences and find new market opportunities. For the last 15 years, she has helped build and launch products and movements across a wide range of sectors and issues for clients like Google, the Center for Global Policy Solutions, and the Rockefeller Foundation.

Before starting Thicket Labs, she worked as a writer, strategist and researcher at venture-backed startups, creative agencies, and market research firms. Deepthi started her career working on state and local policy campaigns in California. Insights from Thicket Labs have been featured in *USA Today*, *Fast Company*, and *Stanford Social Innovation Review*.

thicket

thicketlabs.com

Thicket Labs

The story of our future is determined by the decisions we make today. Thicket Labs is an audience research and innovation company pioneering new analytics, insights, and forecasts for the field.

Since 2014, Thicket Labs has helped clients discover, test, and select ideas and solutions to invest in for the long-term. Clients engage with audiences through Thicket's proprietary gamified learning platform and algorithmic insights.

thicket

thicketlabs.com