

[ABOUT US \(HTTPS://WEB.ARCHIVE.ORG/WEB/20160801033556/HTTP://WWW.THEOCULISTS.COM/ABOUT-US/\)](https://web.archive.org/web/20160801033556/http://www.theoculists.com/about-us/) •

[JOIN US \(HTTPS://WEB.ARCHIVE.ORG/WEB/20160801033556/HTTP://WWW.THEOCULISTS.COM/ABOUT-US/#JOBS\)](https://web.archive.org/web/20160801033556/http://www.theoculists.com/about-us/#jobs)
BY CARL ANDERSON

The Warby Parker Data Book

At Warby Parker, our data dictionary is hugely valuable. Prior to its existence, we had no assurance that analysts on different teams were using the same definitions to define metrics and, even when metrics were defined, we had nowhere to document them. Now, everybody speaks the same language.

Developing this source of truth was a long endeavor. Over a six-month period, our Data Science team rolled out our business intelligence tool, [Looker \(https://web.archive.org/web/20160801033556/http://looker.com/\)](https://web.archive.org/web/20160801033556/http://looker.com/), to our analysts. Part of that process involved implementing the various ways that analysts can slice, dice, and aggregate data. Some dimensions are simple, straightforward, and unambiguous (e.g., group by ZIP code), while others involve domain and business-specific logic (e.g., when one person purchases a gift card and another person redeems it, which of those is considered the customer?). We brought different teams together and facilitated discussions for them to thrash out those definitions and get alignment. The output of those discussions was the canonical definitions of all the metrics and terms used in the company.

Once we had this information, we asked ourselves the following: How do we document this information so that it has good search functionality and user experience? How can everyone access the content? How can we keep it up to date?

Our options

We considered the following options, none of which worked for us:

- **Wiki:** We have a wiki, but there is a large set of content that exists already, and the search functionality is awful; the Data Book would simply get buried.
- **Looker:** At the time, Looker didn't provide a mechanism to document dimensions and measures, although they now provide a mechanism to show plain-text-only definitions on mouseover. It makes sense to tie the documentation to the code in one place, but we wanted company-wide visibility to these definitions—not just Looker user access.
- **GitHub:** We considered markdown documents in GitHub. These could be versioned but offered few other benefits over a wiki. Again, it doesn't provide easy access to all employees; those GitHub pages would have to be public, which we certainly don't want—otherwise, everyone at Warby Parker would need an account, which is more friction and another password to forget.

The solution: GitBook

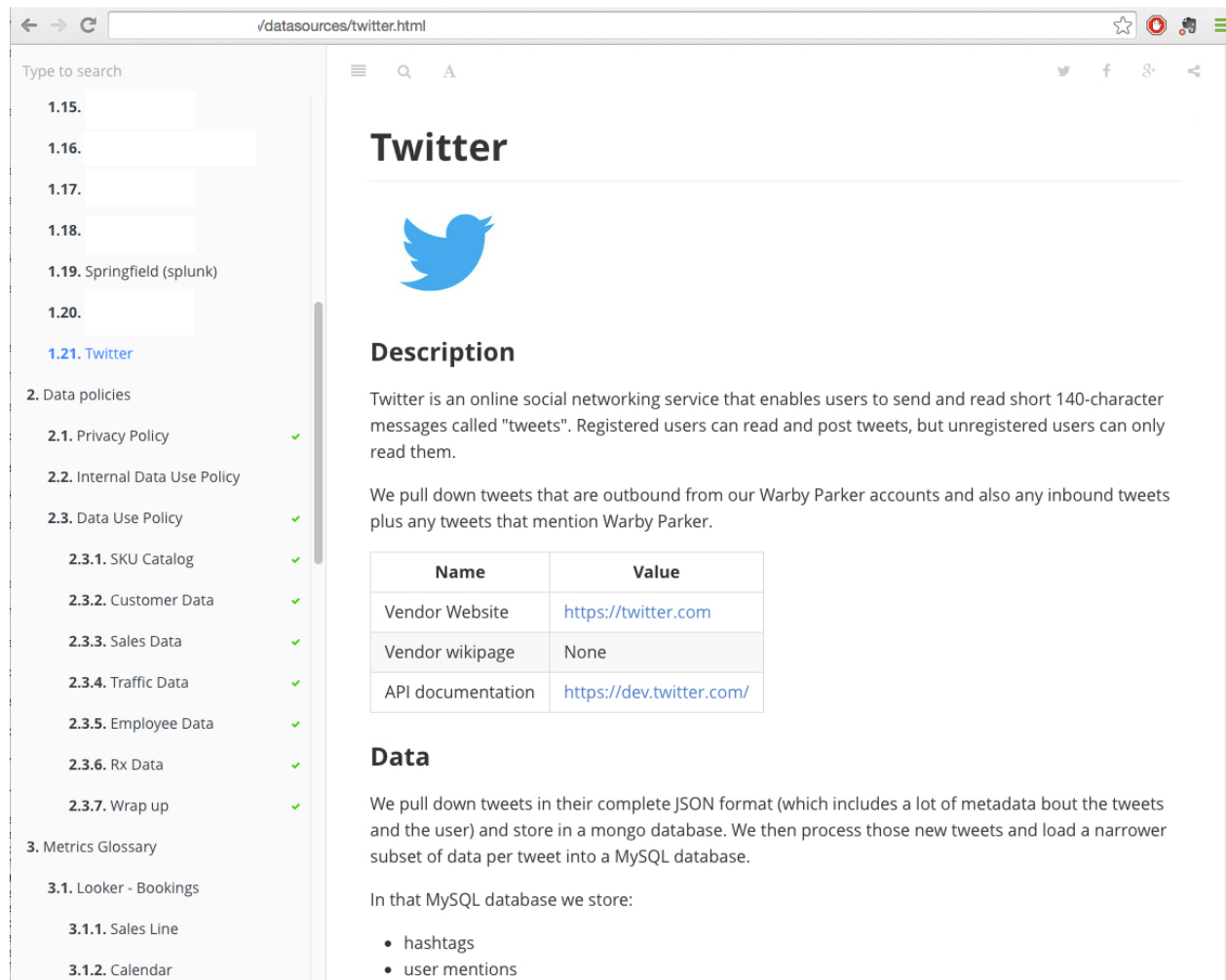
In the end, we opted for [GitBook \(https://web.archive.org/web/20160801033556/https://www.gitbook.com/\)](https://web.archive.org/web/20160801033556/https://www.gitbook.com/). GitBook is a node library that processes pages written in markdown and builds them into an interactive website. It has its own web server, or you can use your own (we use nginx). In addition, a tool called [calibre \(https://web.archive.org/web/20160801033556/http://calibre-ebook.com/help\)](https://web.archive.org/web/20160801033556/http://calibre-ebook.com/help) can convert the content into PDF, .MOBI, or EPUB files. This flexibility makes it easy to consume the Data Book on desktops or mobile devices or to print it as single file.

It met all our needs:

- GitBook provides a beautiful interface. You can page through the content more like a book. It has great search functionality. As you type individual characters of a word, the set of pages narrows down. For instance, by the time you've typed "coo" in the search box, you are shown just the few pages that mention "cookies."
- Under the hood is markdown and HTML (we use HTML for tables, as it provides us more control). That gives us freedom to style and shape the content as we wish.
- It is [open source \(https://web.archive.org/web/20160801033556/https://github.com/GitBookIO/GitBook\)](https://web.archive.org/web/20160801033556/https://github.com/GitBookIO/GitBook), so we can self-host. That means we can provide this as an internal website so that everyone at Warby Parker can access it. The PDF output gives us the option to auto-generate a copy of the content and provide this as part of any new analyst's welcome pack.

With the help of two in-house TechOps employees (one of whom you can see sporting Warby Parker frames in a [previous post](https://web.archive.org/web/20160801033556/http://www.theoculists.com/helping-our-customers-through-text-classification/)), we created a Docker container and ansible script to run the GitBook command to convert the markdown and HTML content to GitBook HTML website content. As part of that script, we also run the calibre command to output the same content as a PDF. We copy both HTML and PDF output to a S3 bucket. With nginx serving up the content, that serverless arrangement is cheaper and works perfectly for this low-traffic site.

We also created a Bamboo plan that monitors the git repo where we version the content. After any content change, the plan runs, then reruns the ansible script and updates the site and PDF. That way, our content and code are as up to date as possible.



(https://web.archive.org/web/20160801033556/http://www.theoculists.com/wp-content/uploads/2015/04/databook_screenshot.jpg).

We love our data book. It allows the whole company to draw from the same source of truth, know that we have a common unambiguous vocabulary, and know that different analyses are comparing apples to apples. The content has expanded to document all our data sources, all our privacy and data use policies, and we will be adding data-related training material. It is now the reference for all things data at Warby Parker. Anyone can access it, and it gets a lot of use. We often hear, "I was going to email you about this metric, but I used the Data Book and found my answer." Thanks GitBook!

Posted on April 2015

< OLDER POST ([HTTPS://WEB.ARCHIVE.ORG/WEB/20160801033556/HTTP://WWW.THEOCULISTS.COM/SYSTEMS-DEVELOPMENT-GAMIFIED/](https://web.archive.org/web/20160801033556/http://www.theoculists.com/systems-development-gamified/))

NEWER POST > ([HTTPS://WEB.ARCHIVE.ORG/WEB/20160801033556/HTTP://WWW.THEOCULISTS.COM/FAST-FREQUENT-AND-FOCUSED-BUILDING-BETTER-USER-TESTING-AT-WARBY-PARKER-2/](https://web.archive.org/web/20160801033556/http://www.theoculists.com/fast-frequent-and-focused-building-better-user-testing-at-warby-parker-2/))

WARBY PARKER (<https://web.archive.org/web/20160801033556/http://www.warbyparker.com/>)

Warby Parker (<https://web.archive.org/web/20160801033556/http://www.warbyparker.com/>)

Buy a Pair, Give a Pair (<https://web.archive.org/web/20160801033556/http://www.warbyparker.com/do-good>) 888.492.7297 (tel:8884927297).

CONNECT WITH US: TWITTER ([HTTPS://WEB.ARCHIVE.ORG/WEB/20160801033556/HTTPS://TWITTER.COM/WARBYPARKER](https://web.archive.org/web/20160801033556/https://twitter.com/warbyparker))
FACEBOOK ([HTTPS://WEB.ARCHIVE.ORG/WEB/20160801033556/HTTPS://WWW.FACEBOOK.COM/WARBYPARKER](https://web.archive.org/web/20160801033556/https://www.facebook.com/warbyparker))

© 2016 THE OCULISTS