Aarhus University
Jakob Munch-Brandt

A dataset from my wardrobe

Characters: 5147

In this report, I will conduct a critical reflection over the process of creating a database based on a personal collection of 20 pieces of clothing from my own wardrobe. Working with transforming physical, personal belongings into digital, structured layout required a big change in perspective. The clothes went from being daily, wearable belongings to data points. I realized quickly that data isn't an objective reflection of reality, but moreof a product constructed through active, curatorial choices. Using Robert Kitchin's Critical Data Studies (2025) and his differentiation of 'data' and 'capta', and Amalie Acker's Metadata chapter (2021), I will analyze the design choices I took that ended up forming the final dataset.

My process started with collecting items from my wardrobe, which was pretty straight forward. I chose 20 pieces of clothing from my closet, with the idea of representing at least a little variation in type and material. Afterwards I began the datafication, which was the most time consuming part. To do this, I created a new excel sheet, and I tried to think of a few different categories of data, such as numerical data, categorical data and image data. I tried to include both categories that depended on my opinion, such as optimal season to wear, but also things like what size it was, found on the label.

I laid every piece of cloth on the table, found the washing label and started to systematically type in the information. The repeating manual proces of looking, reading and typing made me look closer at the things I own, and to illustrate Kitchin's point of capta, I "captured" actively the information that I was looking for to type in the categories I made myself. Other information that wasn't needed to fill in my categories was ignored, such as instructions on how to wash it, even though it was also on the label. Afterwards I digitalized the entire collection by taking pictures of each item. This proces was different and more representative of my mental image of the items, than the data captured. In the end, I had two different, and yet connected, representations of my clothes. One being a spreadsheet of structured data in categories I chose myself, and the other a collection of images.

When I looked at the spreadsheet after, it was interesting to see how my complex, physical items of clothing that I have specific memories with, some still with specific fragrance smells, was suddenly reduced to lines of data in a spreadsheet. Something was lost in the translation, from things like my denim shorts and my soft hoodie both being listed as 100% cotton, and t-shirts fitting me the exact same, but one being S (sold as oversized) and the other being L (sold as slim fit). The result was structured and comparable but also a less true representation of the real thing.

The choices I made when I defined the columns in the spreadsheet was definitely essential for the datasets final form and potential. The spreadsheet contains both data and metadata. The values for each physical object is data, but the image filename and each column in of itself in the spreadsheet are what Acker calls "data about data", also known as metadata. It's interesting to note that you can see which images of the clothing was taken in succession, and which was taking at a later date, or at least where I had taken images inbetween - just from the names.

Aarhus University
Jakob Munch-Brandt

Creating the categories was my own series of subjective choices of what I believed would be important categories. I chose the categories which I believed might be interesting for the report and yet were easy to find on the washing label. I chose primary material as it was a more manageable task than writing all materials, since some were split in 3-4 materials, and I would have needed to create several more categories for that, while some would be without data in those. I did include subjective variables, including the category: “From 1-3, how likely am I to choose this option over other garmant type options in the same category?”. Here I tried to convert a complex, qualitative feeling into a quantitative value. This reduction is a good example of Kitchin's concept of capta. I made a rating scale, but I left our the context. The feedback on the dataset pointed out that richer descriptions could have preserved more nuances, e.g. if a shirt is a favorite because of a specific memory. In the end I deliberately chose the numerical rating to make sure the data was quantifiable. By reducing the feeling to a number it became possible to sort and compare the wardrobe. This underlines the complexity in datafication. To make the world more calculable, we may have to sacrifice some contextual depth.

In the end has this task been a very hands-on and good exercise in the theoretical points of Kitchin and Acker. The process of converting my clothes to data has shown me how creating a dataset isn’t necessarily a neutral mirroring of, for example, my wardrobe, but a deeply personal translation formed by my choices, curiosity and limits. It is very eye opening that the data set isn’t just objective facts, but are formed from subjective choices.

References

Kitchin, R. (2025). Critical data studies: An A to Z guide to concepts and methods. Polity.
Acker, A. (2021). Metadata. In N. B. Thylstrup, D. Agostinho, A. Ring, C. D’Ignazio, & K. Veel (Eds.), Uncertain archives: Critical keywords for big data (pp. 321–330). The MIT Press.