Title: What can online data tell us about eating habits?

Authors: Christoph Trattner and David Elsweiler

Abstract: Understanding how individuals shift to diets with much smaller ecological footprints may help us persuade more people to change their habits and transition to more sustainable food systems. Online interactions provide important answers.

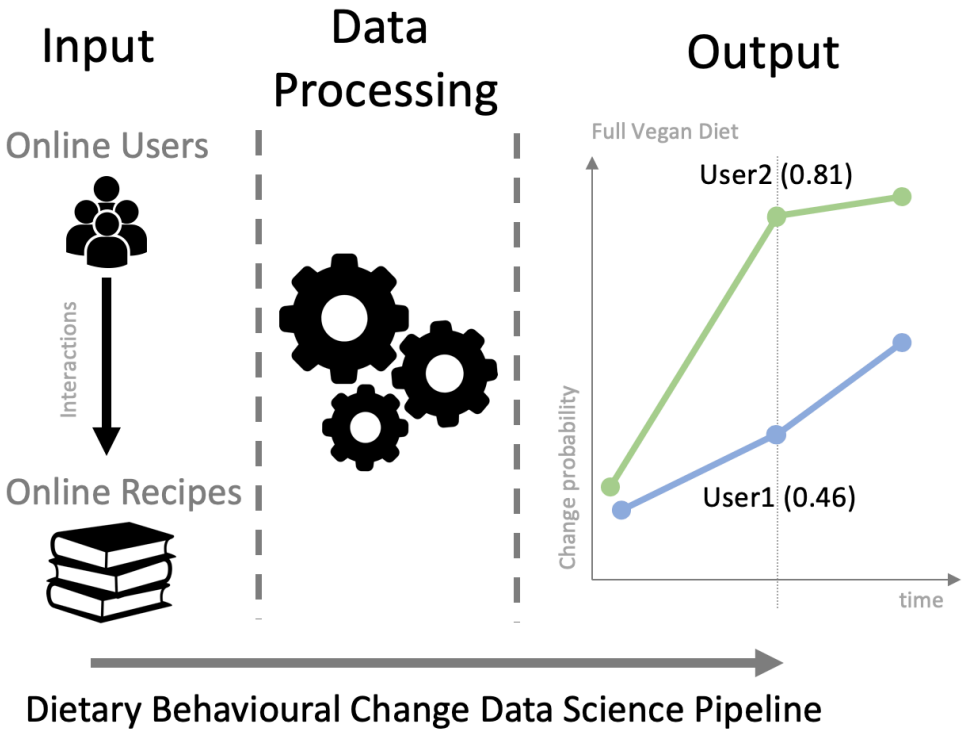

Figure 1: Schema of a data science pipeline (from input to output) to reveal behavioural dietary change from online user-recipe interaction data.

Main Text:

Food's central role for human lives goes beyond mere sustenance and includes health and well-being, social contact and control, and culture and rituals (1). Understanding what humans eat, when and why, has been the research focus of several academic disciplines for decades. Recently a new research domain has emerged, which uses traces of human behaviour -logged via interactions with online recipe portals, search engines and social media- to learn about food preferences and habits. So-called behavioural log data offers several advantages relative to traditional methods. Despite providing rich, detailed insight, conventional methods are often restricted to a small sample of people when using direct observation or, when using surveys, they suffer from the value-action gap -differences between reported and actual preferences.

By analysing such log data, researchers have been able to identify patterns in human behaviour, such as the overall preference for fatty and calorie-rich foods (2), seasonal

trends (3), geographical differences (4) and divergent preferences across genders (5). Using such insights it is possible to algorithmically predict the recipes that will become popular or that individual users will interact with (6). Moreover, online behavioral data have been correlated with prevalence of diet-related illness in the population (7).

These analyses usually aim to understand behaviour to inform systems that can assist people nourish themselves more healthily. Accordingly, these sorts of findings have inspired recipe-suggestion algorithms and adapted user interfaces to influence food choices. Recipe-recommending systems have been shown to 'nudge' users towards healthier options (2). They can also help generate dietary plans that conform to government nutritional guidelines (8).

Writing in Nature Sustainability, Biermann and Asano analyse interaction data from the most popular recipe portal in Germany, Chefkoch.de, comprising 2.5M user interactions with over 240 thousands recipes. This study is exciting for two main reasons. First, the work is motivated on sustainability, rather than on human health and lifestyle-related disease, as most previous studies in this domain. The authors are premised on the idea that eating less meat and animal produce is environmentally friendly.

Second, the analytical approach is novel. Most often, this kind of research is focused on how to change behaviour. However, actual changes in habits are rarely studied. To study such observed behaviour, the authors derive an innovative method that focuses on the tags describing the content and preparation of a recipe. Then they analyse at different levels of granularity the tags applied to recipes.

At a coarse level, they study how users interact over time with recipes featuring specific tags. The authors find a large proportional increase in the sharing of recipes with the tags "vegetarian" and "vegan", while meat-related tags such as "beef", "pork" and "poultry" were not associated with such as rise. They also detect a fast increase in the proportion of recipes submitted to the platform that are vegetarian (30% of the submitted recipes) and vegan (14%). The authors cite survey research to show that these findings mirror current market dynamics in Germany.

In a next step, they investigate the probability of an individual changing their behaviour, such that they interact with recipes associated with tags that they had infrequently interacted with in the past. The authors measure this with a metric referred to as the *infection rate of a given tag*. They moreover derive a metric that measures the extent to which a user continues this pattern of interaction – the tag's *retention rate.* They find complementary evidence for the increasing popularity of meat-free meals. The tags "vegan" and "vegetarian" are associated with a high infection rate whereas the meat-based tags have low and declining infection-rates. Using both infection and retention rates, the authors are able to estimate the frequency of vegetarian and vegan diets from the user population obtaining results close to those reported by large surveys (9).

At the fine level, they explore how individual users transition from meat to meat-free recipes, by analysing the sort of recipes these users interacted most during transitions.

The authors identify strong replacement patterns; while transitioning, the recipes users liked are more similar to the original meat-based recipes than to other recipes. For example, a courgette-patty replacing a meat-based burger-patty.

In a final step, Biermann and Asano complement the study with qualitative data, contacting and interviewing users identified as exhibiting extreme dietary change. The interviews confirm that users prefer to transition by eating meat-free equivalents of favourite meat dishes before they learn about more novel vegetarian dishes as they become more experienced.

In conclusion, this study has two major contributions: an innovative approach to trace dietary change on a large scale using online recipe interaction, and novel findings about how this change looks like. This work opens up new interesting avenues for research. For example, examining temporal dynamics of dietary change can shed light on the duration and drivers of such transitions.

Competing interests: The authors declare no competing interests.

References

1. Anderson, E. N. (2014). *Everyone eats: understanding food and culture*. NYU Press.
2. Elsweiler, D., Trattner, C., & Harvey, M. (2017). Exploiting food choice biases for healthier recipe recommendation. In Proceedings of the 40th international acm sigir conference on research and development in information retrieval (pp. 575-584). ACM.
3. West, R., White, R. W., & Horvitz, E. (2013). From cookies to cooks: Insights on dietary patterns via analysis of web usage logs. In *Proceedings of the 22nd international conference on World Wide Web* (pp. 1399-1410). ACM.
4. Wagner, C., Singer, P., & Strohmaier, M. (2014). The nature and evolution of online food preferences. *EPJ Data Science*, *3*(1), 38.
5. Wagner, C., & Aiello, L. M. (2015). Men eat on Mars, Women on Venus?: An Empirical Study of Food-Images. In WebSci (pp. 63-1).
6. Trattner, C., Moesslang, D., & Elsweiler, D. (2018). On the predictability of the popularity of online recipes. *EPJ Data Science*, *7*(1), 20.
7. Abbar, S., Mejova, Y., & Weber, I. (2015). You tweet what you eat: Studying food consumption through twitter. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems* (pp. 3197-3206). ACM.
8. Trattner, C., & Elsweiler, D. (2017). Investigating the healthiness of internet-sourced recipes: implications for meal planning and recommender systems. In Proceedings of the 26th international conference on world wide web (pp. 489-498). International World Wide Web Conferences Steering Committee.
9. Statista (2018). Vegetarismus und Veganismus. URL: https://de-statista-com.pxz.iubh.de:8443/statistik/studie/id/27956/dokument/vegetarismus-und-veganismus-statista-dossier/