Check for updates

OPEN ACCESS

EDITED BY Shauna Downs, The State University of New Jersey, United States

REVIEWED BY

Carolina Coimbra Vieira, Max Planck Institute for Demographic Research, Germany Marie Laure Charpignon, Massachusetts Institute of Technology, United States

*CORRESPONDENCE Kristina Gligorić ⊠ kristina.gligoric@epfl.ch

RECEIVED 30 May 2023 ACCEPTED 14 May 2024 PUBLISHED 04 June 2024

CITATION

Gligorić K, Zbinden R, Chiolero A, Kıcıman E, White RW, Horvitz E and West R (2024) Measuring and shaping the nutritional environment via food sales logs: case studies of campus-wide food choice and a call to action. *Front. Nutr.* 11:1231070. doi: 10.3389/fnut.2024.1231070

COPYRIGHT

© 2024 Gligorić, Zbinden, Chiolero, Kıcıman, White, Horvitz and West. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

Measuring and shaping the nutritional environment via food sales logs: case studies of campus-wide food choice and a call to action

Kristina Gligorić^{1*}, Robin Zbinden¹, Arnaud Chiolero^{2,3,4}, Emre Kıcıman⁵, Ryen W. White⁵, Eric Horvitz⁶ and Robert West¹

¹Data Science Lab, EPFL, Lausanne, Switzerland, ²Population Health Laboratory (#PopHealthLab), University of Fribourg, Fribourg, Switzerland, ³School of Population and Global Health, McGill University, Montreal, QC, Canada, ⁴Swiss School of Public Health (SSPH+), Zurich, Switzerland, ⁵Microsoft Research, Redmond, WA, United States, ⁶Office of the Chief Scientific Officer, Microsoft, Redmond, WA, United States

Although diets influence health and the environment, measuring and changing nutrition is challenging. Traditional measurement methods face challenges, and designing and conducting behavior-changing interventions is conceptually and logistically complicated. Situated local communities such as university campuses offer unique opportunities to shape the nutritional environment and promote health and sustainability. The present study investigates how passively sensed food purchase logs typically collected as part of regular business operations can be used to monitor and measure on-campus food consumption and understand food choice determinants. First, based on 38 million sales logs collected on a large university campus over eight years, we perform statistical analyses to quantify spatio-temporal determinants of food choice and characterize harmful patterns in dietary behaviors, in a case study of food purchasing at EPFL campus. We identify spatial proximity, food item pairing, and academic schedules (yearly and daily) as important determinants driving the on-campus food choice. The case studies demonstrate the potential of food sales logs for measuring nutrition and highlight the breadth and depth of future possibilities to study individual food-choice determinants. We describe how these insights provide an opportunity for stakeholders, such as campus offices responsible for managing food services, to shape the nutritional environment and improve health and sustainability by designing policies and behavioral interventions. Finally, based on the insights derived through the case study of food purchases at EPFL campus, we identify five future opportunities and offer a call to action for the nutrition research community to contribute to ensuring the health and sustainability of on-campus populations-the very communities to which many researchers belong.

KEYWORDS

food choice, measurement, monitoring, determinants, digital traces, health, sustainability, policy

1 Introduction

The concept of a modern campus evolved to be more than a collection of buildings and grounds that belong to an institution. A modern campus is a complete local community. In such a situated context, people spend significant parts of their time working, educating, or being educated, but also socializing, learning, doing sports, or entertaining themselves.

10.3389/fnut.2024.1231070

People also consume food regularly and globally, in university, corporate, medical, industrial, and other types of campuses. The food consumed on campuses has broad implications for the people on campus and the general population, impacting health and the environment (1, 2). Given the need to actively address the challenges of climate change (3), stakeholders have a growing interest in reducing their campuses' environmental impact.

Social issues, including health and environmental issues, as well as situated socialized contexts such as campus environments, have been areas of interest in nutrition research. The nutrition research community is well-positioned to help campus communities, universities, corporate stakeholders, and policymakers promote these values. To be able to take action toward the goals of ensuring on-campus health and sustainability, it is important to know, to begin with, what foods are consumed on campus and in what context.

However, answering this question poses two major challenges. First, it is challenging to provide good measurements. Traditional methods such as surveys rely on self-reporting, face limitations such as under-reporting of undesired behaviors (4-6) and cannot capture temporal dynamics accurately. Researchers have begun to explore digital traces, e.g., from social media (7), search engines (8-10), or food-tracking applications (11, 12), however, it remains unclear to what extent such distant proxies reflect real food consumption (13-15). Second, measurements from a single campus do not necessarily generalize to other campuses since behaviors are observed at different times, locations, and food landscapes. The worldwide campuses are fundamentally incomparable isolated eco-systems. Hence, much about fundamental campus dietary behaviors with implications for health and the environment-such as meat vs. meat-free meal consumption and consumption of caffeinated or alcoholic drinksremains unknown. New measurement methods are needed to overcome these challenges and collect dynamic, fine-grained, and generalizable data about the diet of campus populations.

In the present work, we investigate how passively sensed¹ food purchase logs can be used to shape the nutritional environment, and argue that nutrition researchers are well-positioned to contribute to improving the nutritional environment at their institutions. To that end, we introduce case studies of an anonymized dataset of passively collected logs of food purchases made on a large university campus and perform statistical analyses of purchasing behaviors with policy implications. The present work studies purchasing behaviors on a university campus, however, the notion of campus, the insights, opportunities, and challenges refer to similar, more or less closed environments.

In this work, we present the following contributions. First, through a case study of purchases at one campus, we report quantitative empirical analyses of passively sensed food purchases. In doing so, we aim to characterize the determinants of food consumption and demonstrate how studies relying on passively sensed purchase logs have the potential to support the stakeholders in their efforts to understand food choice and consumption. We demonstrate the utility of measurement and monitoring via large-scale digital traces through case studies leveraging real-world passively collected food sales logs. In particular, aiming to understand why people on campus purchase the foods they do, we identify spatial proximity, food item pairing, and academic schedules (yearly and daily) as important determinants.

We then formulate a set of further research questions that showcase the breadth of insights that can be sourced from the purchase logs, beyond the specific campus studied here. We aim to make a case for re-purposing such data, which are often available by default and can serve as a valuable source of information to harness in campus environments to monitor and measure nutrition. Finally, we conclude by offering future opportunities and a call to action for the nutrition research community to contribute to ensuring the health and sustainability of the institutions researchers are part of. We argue that it is crucial for researchers to understand the values embedded in the existing food systems at campuses, the existing data collection and analysis practices, and the values and priorities of the individuals impacted by the food offering and food consumption.

2 Materials and methods

2.1 Dataset: food purchase logs

We leverage an anonymized complete dataset of food purchases made on the École Polytechnique Fédérale de Lausanne (EPFL) university campus. The data spans eight years, from 2010 to 2018, and contains about 38 million transactions, of which about 18 million were made with a badge that allows linking to an anonymized person's ID.

The statistical analyses are based on the seven-year period from Jan 1 2012 to Dec 31 2018 with menu data available. The data includes 38.7 k users, who, on median, are observed for a time period spanning 578 days and make 188 transactions. The analyses include all the transactions within that period, unless otherwise noted. Each transaction is attributed with the time it took place, information about the sale location, the purchased items, their quantity, and price. Sold food items are mainly ready to consume and are associated with unstructured textual descriptions (e.g., "coffee", "croissant", "Coca-Cola can").

The data covers all the food outlets permanently located on campus, including restaurants, cafes, and vending machines. During the entirety of the studied period, there were twelve major catered shops located throughout the campus,² and a number of self-service vending machines. The shops typically open Monday through Friday, opening at 07:00 and closing at 20:00, while vending machines are available 24/7. The shops generally are not open during the weekend. All shops offer lunch, while breakfast and dinner offerings vary across shops. Note that student halls in the vicinity of the campus do not have associated dining halls. Hence, the catered shops are the main campus-provided food option for students. In the close proximity of the campus, there are several

¹ Contrary to traditional data sources—e.g., surveys—digital data is passively collected in order to support regular campus operations, meaning that the act of data collection does not impact or interfere with people's behaviors and the measurements do not rely on self-reporting.

² Except for one restaurant, which closed in 2016 and re-opened later in the same year.

non-affiliated food providers, including two seated restaurants (an Italian restaurant and a fast food outlet), two grocery stores, a cafe, and temporarily located food trucks.

In addition to anonymized on-campus purchase logs, we also analyze a smaller-size enriched dataset gathered during a threeweek campus-wide sustainability challenge in November 2018, during which 1,031 consenting and volunteering participants formed 278 teams to compete in taking sustainable actions (e.g., taking the stairs instead of the elevator, or consuming a vegetarian meal). For this subset of individuals, we leverage demographic information and identify all of their transactions present in the purchase logs. This results in a subset of 585,812 transactions annotated with demographic information.

The unstructured food item textual descriptions were additionally mapped to categorical labels such as "sandwich" or "dessert" by a research assistant who labeled the 500 most frequently purchased items, which account for 95.4% of the total volume of item purchases observed in the dataset. Further supplementary information about the dataset are outlined in the Supplementary material.

2.2 Ethical considerations

Nutrition is a potentially sensitive personal behavior. To protect user privacy, the log data used here was accessed exclusively by the personnel involved in this project, and stored and processed exclusively on internal servers. The data was obtained with approval from Data Protection Officer³ and was anonymized before it was made available to the researchers for analysis.

In the data collection and analysis process, we worked directly with campus food-providing administration and transaction system managers who exported and anonymized the data, to understand the information about the food items and restaurants encoded in the dataset. Finally, we note that our work leveraging purchase logs was conducted retrospectively on data that had been collected passively in order to support campus operations. Thus, our analysis did not influence users in any way. Purchases executed by sustainability challenge participants were analyzed retrospectively as well, without influencing them in any way.

3 Results

Motivated by the existing knowledge about human dietary behaviors, we perform descriptive statistical analyses of foodpurchasing behaviors in order to characterize prominent oncampus food choice determinants. Since human dietary behaviors and choices are known to exhibit spatio-temporal regularity (16) and recurrent routine-like patterns (17, 18), we analyze dimensions along which purchasing behaviors vary, including temporal, spatial, and choice regularities (Section 3.1). We then analyze the known issues that campus food consumption faces, including ensuring sustainability and health, and tackling the impacts of regular campus operations (Section 3.2). In what follows, we describe our findings and discuss how these insights can inform on-campus policy-making.

3.1 Spatio-temporal and choice determinants of food purchases

3.1.1 Temporal determinants: the heartbeat of the campus

During a year, the academic calendar dictates life on campus. The three important periods in the academic calendar of the studied campus—semesters, exam sessions, and breaks—are marked in Figure 1. We find that the number of transactions peaks during the spring and fall semesters when students are attending lectures, drops during the exam sessions, and reaches the minimum during the winter and summer breaks. During the course of a day (Figure 2), the transactions peak in the morning, during the time of lunch, and during the afternoon or evening snack time.

3.1.2 Spatial determinants: migration patterns between shops

After temporal, we describe spatial determinants in the purchasing behaviors. We are interested in understanding the regularities in visits to shops. To that end, we perform an association rule analysis. Across all individuals, we consider the shops where a given individual has executed transactions during a week. We then apply the Apriori algorithm (19), an algorithm for the discovery of association rules between shops. Association rules describe regularities between items in transaction data. For example, a rule $\{X\} \rightarrow \{Y\}$ found in the transaction logs would indicate that if a person visited shop X, they are likely to also visit shop Y during the same week. In Figure 3, the graph depicts the confidence of the association rules found using the Apriori algorithm, defined as the percentage of all transactions satisfying Xthat also satisfy Y. This approach lets us monitor shop migration patterns (Figure 3). Rules with confidence greater than the 0.2 threshold are displayed with an edge.

For instance, a thick arrow from node *Shop 10* to node *Shop 4* means that there is a high probability that an individual who went to *Shop 10* also went to *Shop 4* during the same week (shop names are anonymized). *Shop 9* is the on-campus bar that also serves food, while the other shops are on-campus cafeterias that serve meals, beverages, and snacks.

Overall, we find that the distribution of the edges is linked to the geographic locations of the shops. For instance, *Shop 4* is the central place in the graph as there are many arrows with high confidence coming to it, and this cafeteria is indeed at the geographical center of the campus. Moreover, the nearby cafeterias on the campus, frequently visited by students, form a cluster in the shop co-occurrence graph. *Shop 8, Shop 11, Shop 1, Shop 9* and *Shop 2* are near-by cafeterias on the campus, frequently visited by students, and we observe that they form a cluster in the shop co-occurrence graph.

³ The university-appointed Data Protection Officer monitors compliance with data protection laws, informs and advises the university community of their obligations under the law, and assists with issues related to personal data protection.



FIGURE 1

Across the month of the year (on the x-axis), the number of executed transactions (on the y-axis), across the years. Background color marks the yearly academic schedule. Note that in 2017 and 2018, the number of transactions during the semester decreased (dashed lines), likely due to the opening of shops close to campus that do not support purchases with the identifying badge.



3.1.3 Food-choice regularities and routines: co-purchasing patterns

Next, we study the regularities in purchases of specific foods, performing association analysis, as explained above. We consider purchases composed of food items belonging to different categories, i.e., for each purchase, the food categories a person has purchased. In Figure 4, we show the lifts of the association rules found using the Apriori algorithm, defined as the ratio of how frequently the rule appears in the dataset to that expected if *X* and *Y* were independent. Note that the lift is a symmetric measure; therefore, the graph is not directed. We obtain a graph where neighbor nodes are food categories that tend to be purchased together, e.g., *Salty Pastry* and *Pastry. Soup, fruit*, and *salad* are often co-purchased. Note that the *Non-Food* category contains only products linked to beer, such as glasses on deposit at the on-campus bar.

3.1.4 Implications for policies

To summarize, we reveal regularities in how persons oncampus transition between cafeterias during a week and how food items belonging to different food categories are combined. Since the transitions between cafeterias are governed by geographical proximity (Section 3.1), our analyses imply that the decisions about offering at cafeterias should not be taken in isolation, but should account for such spatial migration patterns. Similarly (cf. Section 3.1), the offer should take into account the frequent item pairings because it is not enough to consider foods in isolation (e.g., a sandwich can have good nutritional properties, but if often purchased together with a beverage with high sugar content, promoting it might not be optimal). Analyses of purchase logs can detect these issues and inform measures taken to improve the experiences around food on campus.





3.2 Specific dietary behaviors: sustainability, health, and the academic schedule

3.2.1 Food sustainability: vegetarian meals

We now turn to analyze specific dietary behaviors relevant to sustainability and health on campus. We first analyze vegetarian meal purchases since the consumption and production of meat have a negative impact on the environment (20). Considering the frequency at which vegetarian meals are purchased out of all purchased meals, we find that, on average, across individuals, 5.3% of purchased meals are vegetarian. Vegetarian meals are the most popular among Ph.D. students (11.2%), followed by undergraduate and master students (8.3%), staff (5.9%), and other statuses (e.g., interns and visitors; 2.4%). Vegetarian meals are the most popular among 21–30-year-olds vs. other age groups (16– 20-year-olds and over 30-year-old) and among women vs. men (8.5% vs. 7.2%). We also note that the proportion of vegetarian meals is monotonically increasing over time, rising from less than 3.98% in 2012 to close to 7.96% in 2018. This rise is likely due to the rise in awareness regarding the effect of meat production and consumption on the environment and health, but also due to the university adapting its offering to cater to these trends. The complete distribution is presented in the Supplementary material.

3.2.2 Potentially harmful dietary behaviors

We next analyze purchases of four types of products that could potentially imply harmful effects on health, albeit to a varying extent depending on the number of daily servings and the specific context (including specific nutritional values, ingredients, and ways of preparation, and added sugar content). We focus on (1) beer, (2) energy drinks, (3) coffee, and (4) vending machine items. We measure the fraction of transactions including such products, out of all purchased products, across subpopulations of status, gender, and age.

First, regarding beers, we find that, on average, across individuals, 2.9% of transactions contain a beer. Beer purchases are the most prevalent among "other" statuses, students, Ph.D. students, 21-to-30-year-olds, and men (3.7% men vs. 1.9% women). Monitoring beer purchases is important since consumption of alcoholic beverages in excessive amounts is not recommended (21). Second, on average, across individuals, 0.15% of transactions contain an energy drink. Energy drinks are the most prevalent among students and Ph.D. students, 26-to-30-year-olds, and men (0.1% men vs. 0.041% women). Monitoring energy drinks is pressing since excessive consumption of caffeinated energy drinks have been reported in association with adverse health effects (22). Third, monitoring coffee purchases, on average, across individuals, we find that 15.4% of transactions contain a coffee. Coffee is the most prevalent among women, staff members, and older subpopulations (31-year-olds and older). We note that the question of the effects of drinking coffee on health is nuanced and multifaceted (23). Lastly, monitoring vending machine purchases, on average, across individuals, we find that 6.2% of transactions contain a vending machine item. Vending machine items are the most prevalent among students and 21 to 25-year-olds. Food products available in vending machines often have a high amount of sugar, and vending machines tend to be nutritionally poor (24), highlighting the importance of minitoring such purchases. The complete distribution is presented in the Supplementary material.

3.2.3 Implications for policies

We observe significant differences between subpopulations. Overall, we find that purchases reflecting potentially harmful dietary behaviors are relatively prevalent, especially vending machine purchases (overall, 6.2% or 1 in 16 transactions a person makes contains a vending machine item). Students, Ph.D. students, younger subpopulations, and men are the most susceptible to purchasing potentially harmful items, while vegetarian meals are most popular among Ph.D. students, people between 21 and 30 years of age, and women. These insights can help stakeholders design targeted interventions and campaigns. For instance, campaigns aiming to promote purchases of healthier alternatives might be located and phrased such that they



are geared toward students and younger subpopulations who are the most susceptible to purchasing potentially harmful items. Communicating attractiveness and health of food options through tailored advertisements (25), visual cues (26), and plate graphics (27) might be effective too, as can sustainability challenges. Similarly, interventions providing discounts on healthy food items outside of regular opening hours could be effective in reducing purchases of potentially unhealthy items. Finally, more research is needed to systematically identify social gatherings and events on campus that increase the risk of repeated unhealthy behaviors, and design interventions to address them.

3.2.4 Yearly and daily academic schedules determine on-campus food choice

Finally, we aim to understand how campus operations determine purchasing behaviors. First, to understand how purchasing behavior changes during the exam session, we compare the differences in purchases between semesters and exam sessions. Monitoring weeks of the entire studied period, in Figure 5, we find that, during the exam weeks, compared to semester weeks, there is a significant increase in the relative frequency of purchases of energy drinks (+22.4%), coffee (+9.9%), and a decrease in the relative purchasing frequency of beer (-20.0%), pizza (-15.3%), drinks (-6.0%), and fruit (-5.0%). For instance, the fraction of purchased energy drinks, which has the greatest change between exam weeks and semester weeks, peaks during the fall semester exams (peak occurs on the last week of the year when 1% or 1 in 100 purchased items is an energy drink). As expected, these effects are stronger among students than among staff members, who are less affected by the academic calendar (Supplementary Figures S5, S6).

During the fall and spring semesters, each class begins after the first quarter of each hour and ends at the end of the hour, with a break of 15 minutes between two classes. Investigating the impact of the academic calendar on purchasing behaviors within an hour (Figure 6), we observe different behavior depending on the individual's status and whether the transactions are made during the semesters or not. During the spring and fall semesters, students' transactions peek at the 8th minute in the hour, during the 15minute break, and consequently drop (Figure 6, top). Transactions executed by the staff members do not exhibit such a pattern. Ph.D. students are in between staff and students. These differences between staff members and students disappear during the exam sessions and breaks when there are no 15-min breaks (Figure 6, bottom), implying that hourly patterns are indeed linked with the academic calendar since students tend to take advantage of the 15-min break to buy drink or food.

3.2.5 Implications for policies

To summarize, we find that academic schedules determine food consumption on campus, both at the yearly level (lecture season vs. exam season) and the daily level (lectures vs. breaks). On the yearly level, exam sessions are associated with surges in the consumption of both coffee and energy drinks (Figure 5), while on a daily level, the 15-min break between lectures drives food consumption, particularly among students and during lecture weeks. Exams are associated with increases in purchases of potentially unhealthy products, likely due to stress and performance desires. These insights highlight the need for policy-makers to do more to promote student well-being as stress and anxiety levels are elevated among university and college students (28, 29). Modifying the food offer and making exam sessions a better experience for students by encouraging socialization might be promising directions, as socialization might be reduced during exams since beer and pizza purchases decrease, cf. Figure 5. Such community-setting food environment interventions targeting young adults are typically associated with improvements in diets and nutrition outcomes (30). Similarly, insights regarding the 15-min break imply that purchase line congestion might occur. Interventions mitigating the surge in purchases and modifying the offering might be effective, for instance, by opening additional dedicated checkouts with modified healthy snack offerings.

4 Discussion

4.1 Summary of main findings

The presented results imply that approaches leveraging passively sensed anonymized data should be an essential component in efforts to monitor the evolution of food consumption on campus while being aware of complex spatio-temporal determinants and differences between subpopulations. For instance, the finding that the academic schedules determine food consumption on the studied campus at the yearly and daily level, and the fact that exams are associated with significant increases in purchases of potentially unhealthy products, highlight the need for policy-makers to do more to promote student well-being.

These findings have direct implications for developing new methods for population nutrition monitoring, encouraging better eating practices, and optimizing food offerings. By capturing



nearly all on-campus food consumption, the purchase log analysis approach complements survey-based methodologies, which likely under-report (5, 6) stigmatized consumption of unhealthy items.

4.1.1 Limitations

Certain limitations should be kept in mind when interpreting our results. First, the presented case study is limited to a single campus. Behaviors on the studied campus can fundamentally differ from behaviors on other campuses. External validity and generalization are not guaranteed. Future work should determine to what extent behaviors measured on campus reflect behaviors in other on-campus settings, which might differ in their location, climate, the characteristics of the population, and other environmental properties. We also note that self-selected sustainability challenge participants are not a representative sample of the entire campus population (Supplementary material, Section 1.2). Second, on the studied campus, the unstructured food item labels and the information derived from them are incomplete. For instance, cash transactions cannot be mapped to individuals (46.92% of transactions could be mapped to a specific user). Further information about the difference between the entire dataset, nonidentifiable cash or card transactions, and identifiable transactions executed with the badge are listed in Supplementary Table S1. Overall the distributions are similar across the datasets, although systematic differences exist. Similarly, he vegetarian tag was deduced from the name and type of the product and is, therefore, not necessarily always correct. Third, the log data does not directly capture food consumption but provides indirect proxies via purchasing. We also note that the described analyses do not necessarily establish causality. It is reasonable to assume that students and staff indeed consume the food that they purchase. However, one cannot eliminate the possibility of persons borrowing the card, or paying for items consumed by other people.

Future work should determine to what extent behaviors measured on campus reflect consumption of food purchased outside of the campus and off-campus behaviors. For example, people who primarily eat home-cooked or delivered food might present a skewed representation of their diet in the studied cafeteria purchase logs, and the fact that individuals might consume delivered or home-cooked food introduces unobserved variables into our analyses. Meals consumed at restaurants nearby but outside of campus are similarly not captured, while opening and closing of shops close to campus can impact on-campus behaviors. Furthermore, cultural factors and dietary preferences might play a role in determining whether students and staff choose to rely on food options provided on campus to begin with (31, 32). Related to cultural background, we note that the number of unique users in the dataset grew as the staff and student body expanded (by around 40%, from 12.5 k unique users in 2012 to 17.5 k unique users in 2018). This increase coincides with a recorded growth in the international student population (33), highlighting the importance of campus stakeholders adapting food offerings to various cultures and cuisines.

4.1.2 Further research questions

The insights from the statistical analyses inform a wide range of research questions that can be addressed leveraging anonymized passively sensed purchase logs, potentially in conjunction with other data sources. For instance, future work should understand how food choice is determined by the availability of options, and how the lack of food offerings at late hours might be linked with potentially unhealthy patterns, including vending machine purchases. How does food choice evolve over the course of a day and a week? Do people tend to consume healthier foods at the beginning of the day or the week? To what extent is food choice determined by geographical proximity to options? How much work would people be willing to put in to reach better options? How does the opening of unaffiliated premises near campus affect on-campus meal consumption? Similarly, how does weather affect food purchases? Could random variation in weather be leveraged to study spatial proximity? Additional opportunities to tackle potentially harmful dietary behaviors include understanding their onset. Can on-campus alcohol drinking and energy drink consumption be predicted? How does eating with others impact on-campus food choice, and can social challenges lead to more sustainable on-campus behavior?

4.1.3 Further background and scope

We now turn to situate our findings with respect to the existing literature on monitoring and measuring nutrition with digital traces in a campus-wide setting. A rich body of previous work examined the factors that represent barriers and enablers to healthy eating in campus environments (34). The dominant factors include price, value for money, healthfulness, and taste (35, 36). Previous work has studied the specific, potentially harmful, dietary behaviors in on-campus setting, including the consumption of energy drinks (37, 38), alcohol (39, 40), vending machine foods (24, 41), and a failure to reach the nutritional recommendations (42). Oncampus dietary interventions that aim to improve food availability, accessibility, prices, and promotions through policies received considerable attention from researchers (35, 43-45). Accessibility and price are important factors on the studied campus, too. For instance, during the studied period, the average price of a purchased meal was monotonically increasing, and grew by 30.31% between 2012 to 2019 (from 7.62CHF to 9.93CHF). Similarly, the price of a non-meal item grew from 2.54CHF in 2012 to 3.85CHF in 2019, further highlighting the role of affordability of both on- and off-campus options.

However, understanding all the determinants of food consumption and identifying intervention effectiveness is challenging since the key factors, such as availability and attitudes, change over time, interact, and do not necessarily generalize across campuses. Moving forward, there is a need for unified frameworks to better monitor and understand food consumption in campus environments worldwide.

Studying diets through digital traces has been an active area of research (10, 46, 47). Large-scale passively sensed signals have been harnessed in university campus environments to measure factors of well-being (29) and performance (48), pointing toward the feasibility of leveraging behavioral traces for campus-centric applications (49–51).

Nonetheless, while large-scale digital traces are promising for monitoring and modeling nutrition, little is known about how food sales log signals could be used for understanding and shaping on-campus nutritional environments.

4.2 Opportunities and implications: a call to action

Moving forward, we identify knowledge gaps and outline sketches of themes for analyses to be done at other institutions and other campuses to shape the nutritional environments. We identify five specific areas where we call on the nutrition research community to apply their expertise. Below we propose a research agenda and specify how different research paradigms can answer this call to action to improve nutritional environments.

(1) Deriving generalizable insights about on-campus dietary behaviors. Each campus is an independent eco-system, which makes it challenging to derive findings that hold between campuses. Estimates are typically produced at different times and different locations. Furthermore, cultural factors can, in important ways, alter behaviors—for instance, due to varying susceptibility to stress. What behaviors can generalize across campuses around the world? What behaviors are shared between various types of campuses (e.g., educational vs. industrial vs. corporate vs. medical)?

Future efforts might include creating a network of partner institutions that would enable researchers to replicate the same set of analyses and then perform meta-analyses to discover universal behavioral patterns. We envision developing a system to enable processing the anonymized purchase logs (in a pre-determined format) and obtaining and sharing aggregated high-level insights with other campuses. An a priori-designed meta-analysis would then be performed across a cohort of campuses via "megastudies" (52). A unified client-side analysis framework could be built to process the anonymized logs, following templates of studies that can be conducted worldwide in order to answer pre-agreed research questions.

(2) Collecting more detailed on-campus food offering and consumption data. A major challenge of the efforts to study on-campus food consumption is the lack of transparency regarding food offer and consumption. On campuses, complex sets of factors beyond the knowledge and reach of individuals who consume food often determine the availability of options, the nature of collected data and its usage, resulting in a lack of transparency regarding food. Moving forward, there is a clear need for more robust and open policies. This is an opportunity for nutrition researchers to influence local stakeholders at their respective institutions in order to reject the lack of transparency regarding food offering and consumption, and advocate for sustainable and resilient supply chains (53).

Since institutions are responsible for the health of everyone on campus, there is a pressing need to collect more data about offered food at institutions and campuses globally. Future efforts should involve advocating for and collecting rich information about the food items and communicating it to the public. This information includes but is not limited to foods' origin, distance of the food source from the campus (i.e., whether it originates from local farms), ingredients, nutrients, calories, nutritional scores, sustainability metrics (e.g., carbon footprint), preparation methods, and food waste statistics.

(3) Establishing on-campus digital cohorts. Another potential solution addressing the concern about the incompleteness of purchase logs is setting up digital cohorts where students and staff could share their food consumption and other health-related data for research purposes. In particular, the individuals could share information about their dietary intake, calorie intake, and health outcomes (for instance, through electronic health records or well-being surveys). That would allow studying specific conditions closely linked with dietary habits (e.g., diabetes and heart diseases), and behaviors assumed to be linked with food consumption but not well understood. The latter could include, for instance, information about major life events, daily habits, social media usage and web browsing, menstrual cycle, and mental health. Active data collection efforts are required to answer such ambitious research questions with implications for the health of the general population. Such an effort would be a way to address the incomplete nature of purchase logs by capturing dietary intake and health outcomes more completely.

(4) Discovering needs and priorities regarding the nutritional environment. Since campuses worldwide face a lack of unified effort toward health and sustainability, there is a need for a closer examination of structural processes of power at the campuses and a need for a better understanding of factors that slow down current efforts toward health and sustainability. In addition, there is no principled and systematic understanding of what individuals on campuses worldwide want and need, as well as the difficulties and challenges they face. Future efforts should explore experiences and perspectives from marginalized and underrepresented subpopulations including essential campus service staff who consume food but traditionally do not play a role in the decision-making.

To tackle the challenges mentioned above, new qualitative and participatory approaches are needed. A major lesson learned through our case studies is that the researcher analyzing the logs cannot assume the role of a friendly outsider. The researcher needs to talk to stakeholders, understand the events and schedule of the campus, and engage with both food providers and consumers. In our study, we benefited from having the respective information by being embedded into campus. Similarly, there is a need to acknowledge that food offering, the collected data, and the derived insights embody and reproduce the values of those who designed the food offering systems and data collection mechanisms, to begin with (54). However, the values of individuals who are part of the campus should be discovered, not assumed (55), while enabling individuals on campuses to identify their own priorities and make decisions about the food system and about the future use of digital traces they contribute. In the process, the goal is to give a voice to all stakeholders and enable everyone involved to answer questions such as "How is our data being collected? Will the findings be of benefit to us?" (56).

(5) Developing new principles and practices around ethics and privacy. Since behavioral data can be misused and purchases can reveal potentially sensitive information about individuals (57), it is necessary to balance the potential to do good with the potential to harm. At present, corresponding institutional approvals are needed to perform analyses within one campus, guiding researchers through ethical and privacy concerns. However, new challenges emerge when there is a need to replicate the same analyses across several campuses so as to derive generalizable insights about dietary behaviors. Having all the data, which is potentially sensitive, at a single central point is a risk and a liability.

Future efforts can involve designing an application that would allow locally running purchase log analysis scripts with embedded privacy mechanisms. If there is an agreement about data formats, universal processing scripts can be run locally. Previously, decentralized data processing across silos has been deployed in settings where silos corresponded to hospitals processing medical datasets (58), consistent with the paradigm of federated learning, a privacy-enhancing technique that allows institutions to keep control of the data. Furthermore, federated learning can be used with other privacy-enhancing technologies, such as differential privacy (DP), which ensures that one cannot infer additional information about the original data from the aggregated results (59, 60). To facilitate collaboration over sensitive data, an alternative approach is to take a sensitive dataset as input and generate a structurally and statistically similar synthetic dataset with strong privacy guarantees (61, 62). Finally, secure multiparty computation (63) and trusted environments (64) enable parties to jointly compute a function over their inputs while keeping those inputs private from each other, or from the main processor of the central server. A key barrier toward that goal is determining whether other institutions are willing to provide data to their own researchers to begin with, and whether researchers could share aggregated insights with researchers at other institutions.

4.3 Conclusion

On-campus food offer and consumption broadly affect health, performance, and the environment. We make a case for shaping such nutritional environments by leveraging passively sensed food sales logs, typically available by default. Through case studies of food choice determinants in a large university campus, we demonstrate how analyses of such logs can potentially inform policy-making and argue that nutrition researchers are well-positioned to apply the expertise necessary to contribute to improving food offer and consumption across institutions.

Data availability statement

The original contributions presented in the study are included in the article/Supplementary material, further inquiries can be directed to the corresponding author.

Ethics statement

The studies involving humans were approved by EPFL Data Protection Officer (the university-appointed Data Protection Officer monitors compliance with data protection laws, informs and advises the university community of their obligations under the law, and assists with issues related to personal data protection). The studies were conducted in accordance with the local legislation and institutional requirements. Written informed consent for participation was not required from the participants or the participants' legal guardians/next of kin in accordance with the national legislation and institutional requirements.

Author contributions

KG and RZ contributed to data curation, formal analysis, and writing – original draft. AC, EK, RWh, EH, and RWe contributed to conceptualization and methodology. All authors contributed to writing – review and editing.

Funding

The author(s) declare financial support was received for the research, authorship, and/or publication of this article. We acknowledge support from Microsoft (Swiss Joint Research

References

1. Spencer M, Cienfuegos C, Guinard JX. The Flexitarian FlipTM in university dining venues: Student and adult consumer acceptance of mixed dishes in which animal protein has been partially replaced with plant protein. *Food Qual Pref.* (2018) 68:50–63. doi: 10.1016/j.foodqual.2018.02.003

2. Too L, Bajracharya B. Sustainable campus: engaging the community in sustainability. *Int J Sustain Higher Educ.* (2015) 16:1. doi: 10.1108/IJSHE-07-2013-0080

3. Colglazier W. Sustainable development agenda: 2030. *Science*. (2015) 349:aad2333. doi: 10.1126/science.aad2333

4. Gibson RS, Charrondiere UR, Bell W. Measurement errors in dietary assessment using self-reported 24-hour recalls in low-income countries and strategies for their prevention. Adv Nutr. (2017) 8:980–91. doi: 10.3945/an.117.016980

5. Gemming L, Ni Mhurchu C. Dietary under-reporting: what foods and which meals are typically under-reported? *Eur J Clini Nutr.* (2016) 70:640–1. doi: 10.1038/ejcn.2015.204

6. Boniface S, Kneale J, Shelton N. Drinking pattern is more strongly associated with under-reporting of alcohol consumption than socio-demographic factors: evidence from a mixed-methods study. *BMC Public Health.* (2014) 14:1297. doi: 10.1186/1471-2458-14-1297

7. Abbar S, Mejova Y, Weber I. You tweet what you eat: studying food consumption through Twitter. In: *Proc. of the 33rd Conference on Human Factors in Computing Systems (CHI)*. New York, NY: Association for Computing Machinery (2015).

8. West R, White RW, Horvitz E. From cookies to cooks: insights on dietary patterns via analysis of web usage logs. In: *Proc. of the 22nd International Conference on World Wide Web (TheWebConf)*. New York, NY: Association for Computing Machinery (2013).

9. Vosen S, Schmidt T. Forecasting private consumption: survey-based indicators vs. Google trends. J Forecast. (2011) 30:1213. doi: 10.1002/for.1213

10. Gligorić K, Chiolero A, Kiciman E, White RW, West R. Population-scale dietary interests during the COVID-19 pandemic. *Nat Commun.* (2022) 13:1073. doi: 10.1038/s41467-022-28498-z

Center). Open access funding by Swiss Federal Institute of Technology in Lausanne (EPFL).

Conflict of interest

EK, RWh, and EH were employed by Microsoft Research.

The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fnut.2024. 1231070/full#supplementary-material

11. Achananuparp P, Weber I. Extracting food substitutes from food diary via distributional similarity. In: *10th ACM Conference on Recommender Systems*. New York, NY: Association for Computing Machinery (2016).

12. Bin Morshed M, Kulkarni SS, Li R, Saha K, Roper LG, Nachman L, et al. A Real-Time eating detection system for capturing eating moments and triggering ecological momentary assessments to obtain further context: system development and validation study. *JMIR mHealth uHealth*. (2020) 8:e20625. doi: 10.2196/20625

13. Gligorić K, Djordjević I, West R. Biased bytes: on the validity of estimating food consumption from digital traces. In: *Proc of the ACM Conference on Computer Supported Cooperative Work and Social Computing (CSCW)*. New York, NY: Association for Computing Machinery (2022).

14. Lazer D, Hargittai E, Freelon D, Gonzalez-Bailon S, Munger K, Ognyanova K, et al. Meaningful measures of human society in the twenty-first century. *Nature*. (2021) 595:7866. doi: 10.1038/s41586-021-03660-7

15. Olteanu A, Castillo C, Diaz F, Kiciman E. Social data: Biases, methodological pitfalls, and ethical boundaries. *Front Big Data*. (2019) 2:13. doi: 10.3389/fdata.2019.00013

16. Cuttone A, Lehmann S, Gonzalez MC. Understanding predictability and exploration in human mobility. *EPJ Data Sci.* (2018) 7:1–17. doi: 10.1140/epjds/s13688-017-0129-1

17. Jastran MM, Bisogni CA, Sobal J, Blake C, Devine CM. Eating routines. Embedded, value based, modifiable, and reflective. *Appetite*. (2009) 52:127-36. doi: 10.1016/j.appet.2008.09.003

18. Bisogni CA, Jastran M, Blake CE. The construction of eating episodes, food scripts, and food routines. In: *Handbook of Behavior, Food and Nutrition*. New York, NY: Springer Science & Business Media (2011). p. 987–1009.

19. Agrawal R, Mannila H, Srikant R, Toivonen H, Verkamo AI. Fast discovery of association rules. *Adv Knowl Discov Data Mining*. (1996) 12:259–89. doi: 10.1023/A:1009748302351

20. WWF. Bending the Curve: The Restorative Power of Planet-Based Diets. (2020). Available online at: https://www.worldwildlife.org/publications/bending-the-curve-the-restorative-power-of-planet-based-diets; http://www.sciencedirect.com/science/article/pii/S0306437907000506 (accessed May 23, 2024).

21. Sohrabvandi S, Mortazavian AM, Rezaei K. Health-related aspects of beer: a review. *Int J Food Propert*. (2012) 15:487627. doi: 10.1080/10942912.2010.487627

22. Seifert SM, Schaechter JL, Hershorin ER, Lipshultz SE. Health effects of energy drinks on children, adolescents, and young adults. *Pediatrics*. (2011) 127:3592. doi: 10.1542/peds.2009-3592

23. Poole R, Kennedy OJ, Roderick P, Fallowfield JA, Hayes PC, Parkes J. Coffee consumption and health: Umbrella review of meta-analyses of multiple health outcomes. *BMJ*. (2017) 359:j5024. doi: 10.1136/bmj.j5024

24. Grech A, Hebden L, Roy R, Allman-Farinelli M. Are products sold in university vending machines nutritionally poor? A food environment audit. *Nutr Dietet*. (2017) 74:12332. doi: 10.1111/1747-0080.12332

25. Mathur MB, Peacock J, Reichling DB, Nadler J, Bain PA, Gardner CD, et al. Interventions to reduce meat consumption by appealing to animal welfare: meta-analysis and evidence-based recommendations. *Appetite*. (2021) 164:105277. doi: 10.1016/j.appet.2021.105277

26. Mathur MB, Peacock JR, Robinson TN, Gardner CD. Effectiveness of a theory-informed documentary to reduce consumption of meat and animal products: three randomized controlled experiments. *Nutrients*. (2021) 13:4555. doi: 10.3390/nu13124555

27. Marcone AL, Darmstadt GL, Challamel GA, Mathur MB, Gardner CD. Effects of an educational planetary plate graphic on meat consumption in a Stanford University dining hall: a randomized controlled trial. *BMC Nutr.* (2023) 9:106. doi: 10.1186/s40795-023-00764-3

28. Haidar SA, De Vries N, Karavetian M, El-Rassi R. Stress, anxiety, and weight gain among university and college students: a systematic review. *J Acad Nutr Dietet*. (2018) 118:15. doi: 10.1016/j.jand.2017.10.015

29. Wang R, Chen F, Chen Z, Li T, Harari G, Tignor S, et al. StudentLife: assessing mental health, academic performance and behavioral trends of college students using smartphones. In: *Proceedings of the 2014 ACM International Joint Conference on Pervasive and Ubiquitous Computing*. New York, NY: Association for Computing Machinery (2014).

30. Downs S, Demmler KM. Food environment interventions targeting children and adolescents: a scoping review. *Global Food Secu.* (2020) 27:100403. doi: 10.1016/j.gfs.2020.100403

31. Vieira CC, Lohmann S, Zagheni E, Vaz de Melo PO, Benevenuto F, Ribeiro FN. The interplay of migration and cultural similarity between countries: evidence from Facebook data on food and drink interests. *PLoS ONE.* (2022) 17:e0262947. doi: 10.1371/journal.pone.0262947

32. Obradovich N, Özak Ö, Martin I, Ortuño-Ortín I, Awad E, Cebrián M, et al. Expanding the measurement of culture with a sample of two billion humans. *J Royal Soc Interface*. (2022) 19:20220085. doi: 10.1098/rsif.2022. 0085

33. Riaño Y, Lombard A, Piguet E. How to explain migration policy openness in times of closure? The case of international students in Switzerland. *Globalisat Soc Educ.* (2018) 16:295-307. doi: 10.1080/14767724.2017. 1412823

34. Sogari G, Velez-Argumedo C, Gómez MI, Mora C. College students and eating habits: A study using an ecological model for healthy behavior. *Nutrients.* (2018) 10:1823. doi: 10.3390/nu10121823

35. Roy R, Soo D, Conroy D, Wall CR, Swinburn B. Exploring university food environment and on-campus food purchasing behaviors, preferences, and opinions. J Nutr Educ Behav. (2019) 51:3. doi: 10.1016/j.jneb.2019. 03.003

36. Block JP, Gillman MW, Linakis SK, Goldman RE. "If it tastes good, I'm drinking it": qualitative study of beverage consumption among college students. *J Adolesc Health*. (2013) 52:17. doi: 10.1016/j.jadohealth.2012.11.017

37. Bulut B, Beyhun NE, Topbaş M, Can G. Energy drink use in university students and associated factors. *J Community Health.* (2014) 39:1004–11. doi: 10.1007/s10900-014-9849-3

38. Spierer DK, Blanding N, Santella A. Energy drink consumption and associated health behaviors among university students in an urban setting. *J Community Health.* (2014) 39:132–8. doi: 10.1007/s10900-013-9749-y

39. Davoren MP, Demant J, Shiely F, Perry IJ. Alcohol consumption among university students in Ireland and the United Kingdom from 2002 to 2014: a systematic review. *BMC Public Health.* (2016) 16:173. doi: 10.1186/s12889-016-2843-1

40. Foxcroft DR, Moreira MT, Santimano NMA, Smith LA. Social norms information for alcohol misuse in university and college students. *Cochrane Database Syst Rev.* (2015) 2015:CD006748. doi: 10.1002/14651858.CD006748.pub4

41. Viana J, Leonard SA, Kitay B, Ansel D, Angelis P, Slusser W. Healthier vending machines in a university setting: effective and financially sustainable. *Appetite*. (2018) 121:263–7. doi: 10.1016/j.appet.2017.11.094

42. Strawson C, Bell R, Downs S, Farmer A, Olstad D, Willows N. Dietary patterns of female university students: with nutrition education. *Can J Diet Pract Res.* (2013) 74:138–42. doi: 10.3148/74.3.2013.138

43. Roy R, Rangan A, Hebden L, Louie JCY, Tang LM, Kay J, et al. Dietary contribution of foods and beverages sold within a university campus and its effect on diet quality of young adults. *Nutrition*. (2017) 34:13. doi: 10.1016/j.nut.2016.09.013

44. Deliens T, Van Crombruggen R, Verbruggen S, De Bourdeaudhuij I, Deforche B, Clarys P. Dietary interventions among university students: a systematic review. *Appetite.* (2016) 105:14–26. doi: 10.1016/j.appet.2016.05.003

45. Boronowsky RD, Zhang AW, Stecher C, Presley K, Mathur MB, Cleveland DA, et al. Plant-based default nudges effectively increase the sustainability of catered meals on college campuses: three randomized controlled trials. *Front Sustain Food Syst.* (2022) 6:1001157. doi: 10.3389/fsufs.2022.1001157

46. De Choudhury M, Sharma S, Kiciman E. Characterizing dietary choices, nutrition, and language in food deserts via social media. In: *Proceedings of the 2016 ACM Conference on Computer Supported Cooperative Work and Social Computing* (CSCW). New York, NY: Association for Computing Machinery (2016).

47. Cordeiro F, Epstein DA, Thomaz E, Bales E, Jagannathan AK, Abowd GD, et al. Barriers and negative nudges: exploring challenges in food journaling. In: *Proceedings* of the 2015 Conference on Human Factors in Computing Systems (CHI). New York, NY: Association for Computing Machinery (2015).

48. Wang R, Harari G, Hao P, Zhou X, Campbell AT. SmartGPA: how smartphones can assess and predict academic performance of college students. In: *Proceedings of the 2015 ACM International Joint Conference on Pervasive and Ubiquitous Computing*. New York, NY: Association for Computing Machinery (2015).

49. Morshed MB, Kulkarni SS, Saha K, Li R, Roper LG, Nachman L, et al. Food, Mood, Context: Examining College Students' Eating Context and Mental Well-Being. *ACM Trans Comp Healthc.* (2022) 3:1–26. doi: 10.1145/3533390

50. Saha K, Yousuf A, Boyd RL, Pennebaker JW, De Choudhury M. Social media discussions predict mental health consultations on college campuses. *Sci Rep.* (2022) 12:123. doi: 10.1038/s41598-021-03423-4

51. Swain VD, Kwon H, Saket B, Morshed MB, Tran K, Patel D, et al. Leveraging WiFi network logs to infer student collocation and its relationship with academic performance. *EPJ Data Sci.* (2023) 12:22. doi: 10.1140/epjds/s13688-023-00398-2

52. Milkman KL, Gromet D, Ho H, Kay JS, Lee TW, Pandiloski P, et al. Megastudies improve the impact of applied behavioural science. *Nature*. (2021) 600:478-83. doi: 10.1038/s41586-021-04128-4

53. Davis KF, Downs S, Gephart JA. Towards food supply chain resilience to environmental shocks. *Nature Food*. (2021) 2:54-65. doi: 10.1038/s43016-020-00196-3

54. Friedman B. Value-sensitive design. Interactions. (1996) 3:242493. doi: 10.1145/242485.242493

55. Le Dantec CA, Poole ES, Wyche SP. Values as lived experience: Evolving value sensitive design in support of value discovery. In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI)*. New York, NY: Association for Computing Machinery (2009).

56. Ghoshal S, Bruckman A. The role of social computing technologies in grassroots movement building. *ACM Trans Comp-Human Interact.* (2019) 26:3318140. doi: 10.1145/3318140

57. Hill K. How Target Figured Out a Teen Girl was Pregnant Before her Father Did. Jersey City, NJ: Forbes, Inc. (2012).

58. Andreux M, Terrail JOd, Beguier C, Tramel EW. Siloed federated learning for multi-centric histopathology datasets. In: *Domain Adaptation and Representation Transfer, and Distributed and Collaborative Learning*. New York, NY: Springer Science & Business Media (2020).

59. Dwork C, McSherry F, Nissim K, Smith A. Calibrating noise to sensitivity in private data analysis. In: *Theory of Cryptography Conference*. Cham: Springer (2006).

60. Dwork C, Roth A, et al. The algorithmic foundations of differential privacy. Found Trends Theoret Comp Sci. (2014) 9:211–407. doi: 10.1561/0400000042

61. Ping H, Stoyanovich J, Howe B. Datasynthesizer: privacy-preserving synthetic datasets. In: *Proceedings of the 29th International Conference on Scientific and Statistical Database Management*. New York, NY: Association for Computing Machinery (2017).

62. Jordon J, Yoon J, Van Der Schaar M. PATE-GAN: generating synthetic data with differential privacy guarantees. In: *International Conference on Learning Representations*. (2018).

63. Goldreich O. Secure Multi-Party Computation. (1998).

64. Sabt M, Achemlal M, Bouabdallah A. Trusted execution environment: what it is, and what it is not. In: 2015 IEEE Trustcom/BigDataSE/ISPA. Helsinki: IEEE. (2015).