#Healthy #Fondue #Dinner: Analysis and Inference of Food and Drink Consumption Patterns on Instagram

Thanh-Trung Phan
Idiap Research Institute and EPFL
Switzerland
tphan@idiap.ch

Daniel Gatica-Perez
Idiap Research Institute and EPFL
Switzerland
gatica@idiap.ch

ABSTRACT
Social media generate large-scale data to study food and drink consumption in everyday life. Using Instagram posts in Switzerland over five years, our goal is two-fold. First, we extract key food & drink consumption patterns, through the lenses of a data-driven dictionary of popular items extracted from hashtags, and of a food categorization system used by the Swiss Federal government for national statistics purposes. Patterns related to spatial and temporal distributions of food & drink consumption, demographics, and eating events are extracted and compared to official statistics. Second, using the insights from this analysis, we define two eating event classification tasks, including a two-class task (healthy vs. unhealthy) and a six-class task (the three main meals breakfast/lunch/dinner/ plus brunch/coffee/tea). Both tasks use hashtags as labels for supervised learning. We study how content (hashtags and food categories), context (time and location), and social features (likes) can discriminate these eating events. A random forest and a combination of content and context features can classify healthy vs. unhealthy eating posts with 85.8% accuracy, and the six daily eating occasions with 61.7% accuracy.

ACM Classification Keywords
H.5.m. Information Interfaces and Presentation (e.g. HCI): Miscellaneous

Author Keywords
Instagram; Foursquare; Hashtags; Food; Drink; Consumption Patterns; Human-centered Computing; Social Media.

1. INTRODUCTION
Studying patterns of food and drink (F&D) consumption has been a research subject in academia, government, and the food industry for years. In the past, researchers and governments largely relied on data collected offline [1], like retrospective surveys and phone interviews, which are not easy to obtain and are subject to recall biases and other issues. The ubiquitous use of smartphones and social media has generated new large-scale data sources through which food consumption in everyday life can be studied. This is the case for photo and location sharing services like Instagram (300 million monthly-active users worldwide) and Foursquare (8 billion check-ins worldwide), which opens new, data-driven questions for mobile and ubiquitous multimedia research.

As people share their real-time activities, mood, photos, and location at particular venues, certain aspects of the experience of eating and drinking in daily life can be examined, e.g. how people get benefits from sharing food information on social media and what challenges they encounter [11]. Instagram data (photos and video with captions, hashtags, user mentions, likes, and comments) can be enriched with Foursquare venue data (venues name, category, and additional metadata) through check-ins, which results in a rich description of the context in which F&D information is posted.

To investigate food-related phenomena on mobile social media, extracting and categorizing food and non-food content from posts is a first necessary step. Clearly, users posting photos at a restaurant do not necessarily imply that such photos depict food; they could be photos of people or views of the venue [33]. In addition, food posts often contain text content like hashtags and captions that describe the photo content but also the user location and activities. Hashtags are an essential data source to study the characteristics of food as well as the context and interests of users related to food [19, 15, 5, 34, 31].

In public health and nutrition science studies, the national level is often the fundamental target [7, 9, 14]. This is important because, despite globalization, eating still depends on the local context, i.e., the available food items and brands, food stores, and cooking styles are often local or regional. Furthermore, government statistics, which are essential to ground some of the findings from social media analytics, are most often available nationally. Much of the work on social media and food has focused on English-speaking countries (US and UK) [19, 15, 5, 34], or has discussed global trends only scratching the surface with respect to the existing scientific knowledge about eating patterns in specific countries [31]. We focus our analysis on Switzerland, a multilingual European country for which there is government data on food and eating [1, 3, 4], and in which Instagram is popular among youth. Our work thus adds diversity to the countries studied under the social media and food angle.

MUM 2017, November 26–29, 2017, Stuttgart, Germany

ISBN 978-1-4503-5378-6/17/11. . . $15.00

DOI: https://doi.org/10.1145/3152832.3152857

In this paper, we investigate two research questions:

RQ1: What food and drink consumption patterns can be extracted from geo-localized Instagram/Foursquare Swiss data, and how do these patterns compare to other sources of national statistical data?

RQ2: How do content and context-related features of Instagram food posts can be used to automatically classify everyday self-reported events, like eating healthy or having lunch?

To obtain answers to these questions, we filter all F&D posts from a pool of 2.8 million Instagram posts in Switzerland over a period of five years. Posts contain images, captions, hashtags, comments, timestamp, venue information, and other metadata. For our analysis, we first define a hashtag-based dictionary of F&D (both food items and food-related concepts), based on their specific popularity of the country under study. We then extract general patterns of F&D posting in terms of time, location, and types of eating events. We deepen the analysis by mapping the F&D dictionary into categories defined by the Swiss Federal Food Safety and Veterinary Office (FFSVO) [4]. This allows for a more systematic analysis of the F&D consumption reported on Instagram, including a gender-based comparison. Whenever possible, the patterns extracted from Instagram/Foursquare are compared with figures from the Swiss Federal Statistical Office (FSO) [1] and with menuCH, the first national survey on food consumption in Switzerland conducted over 2014-2015, which used a combination of pencil-and-paper questionnaires with face-to-face interviews with trained dieticians in ten centers throughout the country [3]. Some of the biases of Instagram data appear evident through these comparisons with traditional instruments for collection of food consumption data.

Finally, using the insights from the descriptive analysis, we define automatic inference tasks for two ways of conceptualizing eating events, namely a two-class task (healthy vs. unhealthy eating) and a six-class task (breakfast/brunch/lunch/dinner/coffeetime/teatime). Both tasks use self-reported labels (in the form of hashtags) for supervised learning. Our goal is to understand how content (hashtags and food categories), context (time and location), and social features (likes) can be informative of different eating events. A random forest approach shows that healthy vs. unhealthy eating posts can be inferred with 85.8% accuracy (with content features as most relevant), while the six daily eating occasions can be correctly inferred with 61.7% accuracy (with context features as most relevant).

The rest of the paper is structured as follows. Section 2 discusses related work. Section 3 describes the datasets used in our study. Section 4 describes the generation of the hashtag-based F&D dictionary. Section 5 presents the analysis of spatio-temporal patterns and eating events. Section 6 presents the analysis based on FFSVO food categories. Section 7 presents the automatic inference tasks and discusses the experimental results. Section 8 concludes the paper.

2. RELATED WORK

In this section, we review work related to food post recognition and analysis of food consumption patterns in social media.

Food Post Recognition. Methods to recognize specific food items in social media involve text and images. On Twitter, several works have analyzed text content like hashtags and key terms to recognize tweets containing F&D [5, 19, 15]. In other text-based research, works have investigated food items by examining the textual content and distinguishing the presentation of high-calorie, low-nutrient food items vs. fruits and vegetables [17], and by combining text topics with nutritional fact [20]. Other works have processed Instagram hashtags to detect canonical names and retrieve nutrition information from online sources [34]. Hashtags have the advantages of simplicity and direct semantics, but also have limitations due to polysemy, so it might be hard to identify some food items: for example, *orange* can be both food and a color.

The photos available in Instagram can be of great help to complement hashtags. Recent work [31] has used a hashtag-driven approach to discover the most popular food categories in a given Instagram dataset, and to learn visual recognizers of food images via supervised learning. This work inspires us to extract a data-driven dictionary of F&D items for the specific country under study, which will reflect national trends for food items.

Deep learning has become the preferred choice for learning visual food item classifiers [18, 6], and has been applied on Flickr data for food and non-food classification [30], Instagram [31], and other sources of online data connected to cooking recipes [10], and restaurant menus [27], often with the ultimate interest of counting calories by recognizing the contents of a single image and then extracting nutritional content such as calories. In our work, we do not aim to recognize food images automatically, but rather to use all other available information (hashtags, place and temporal context, and social features) to discriminate types of eating occasions, as opposed to identify food items or caloric content.

Analysis of Food Consumption Patterns. In the social media literature, various F&D consumption patterns have been studied. Alcohol drinking tweets throughout regions across the UK were tracked in [19]. A set of 27 health-related statistics of Twitter data at the US county level, including a few eating-related patterns (limited healthy food, fast food, diabetes, obesity) was studied in [13]. Another study of food-related tweets discovered correlations with obesity and diabetes rates at the US county level [5]. Using data from Instagram, further connections between food-related hashtags (#foodporn and others) and geographic aggregates of obesity indicators in the US context were studied in [26]. As we discussed in the next section, data-driven, food-related dictionaries built from popular hashtags [5, 26] can have large variations across countries. This highlights the need to understand (and be sensitive to) these national differences.

Other work [35] has looked at differences of reported F&D habits through check-ins in Foursquare [24], filtering check-ins at Food and Nightlife venues and allocating them to three classes: drink, fast food, and slow food. In a larger context, all this work is related to the interest on discovering links between geo-localized social media posts and socioeconomic characteristics of local people [22]. Our work focuses on extracting patterns of F&D consumption within a particular country as reflected on Instagram, which is not representative of the full population yet corresponds to a young population as we discussed later in the paper.

Finally, a recent study with Instagram users [11] interviewed 16 women who posted about food to support themselves and others to maintain healthy eating behaviors, through the use of hashtags used for food tracking like #fooddiary, #foodjournal, and #caloriecounting. Inspired by this qualitative research, in our work we studied the feasibility of recognizing self-reported healthy or otherwise eating occasions, indicated by the use of hashtags like #healthyfood, through the use of text, context, and social features.

3. DATASETS

Initially, we defined a spatial grid covering Switzerland (each square in the grid was 111×111 m^2), and use the Instagram API to scan all available venues in the country, for a total of 183K venues. From these venues, we downloaded 2.88 million photos, along with captions and metadata, posted by 594K distinct users between October 2010 and April 2016. In this work, we focus on posts with at least one hashtag between November 1, 2010 and March 31, 2016. This resulted in 1.7 million posts. In the rest of the paper, we call this dataset the *Instagram 1.7M dataset*.

At the time of data collection, Instagram supported matching Instagram venues to Foursquare (4sq) venues, which have richer information such as venue categories (by sending 4sq venue IDs to Instagram, the corresponding Instagram venue were returned). To achieve this, we first downloaded 169K 4sq venues in Switzerland by using Foursquare Venue API Endpoint[1]. After matching the 4sq and Instagram venues, we obtained a total of 84K matched venues. Each venue belongs to a category tree declared by 4sq[2]. This tree has many levels, and each level has a list of category nodes. In this work, we only focus on the top category for each venue. There are ten top categories declared by 4sq: None (-1), Arts & Entertainment (0), College & University (1), Events (2), Food (3), Nightlife Spots (4), Outdoors & Recreation (5), Professional & Other Places (6), Residence (7), Shop & Services (8), and Travel & Transport (9).

In order to define a hashtag-based Food & Drink vocabulary for the specific Swiss case, we started by examining all Instagram posts generated at venues that specifically match 4sq food venues, i.e., category (3) above. The assumption is that posts generated in food venues probably contain more hashtags involving Food & Drink than other venues. This step was practically important as F&D is only one of the hundreds of topics talked about on Instagram. Following this step, we obtained 3,745 matched food venues between Instagram and 4sq. From these food venues, we harvested a set of 65K Instagram posts. As described in detail in the next Section, this data set was used to define our data-driven, hashtag-based dictionary of F&D items, which consists of 184 items. We call this dataset the *Instagram 65K dataset*.

[1]https://developer.foursquare.com/docs/venues/venues
[2]https://developer.foursquare.com/categorytree

Properties	95K Dataset	55K Dataset
# of images	95K	55K
# of total hashtags	1M	576K
# of unique hashtags	136K	80K
# of users	42K	26K
# of venues	22K	10K

Table 1. Instagram datasets used in the rest of the paper.

With the 184-item F&D dictionary, we revisit the larger *Instagram 1.7M dataset* to extract all posts containing at least one F&D item in the dictionary, i.e., harvesting as many images as possible that use our data-driven F&D dictionary. As a result, we obtained 95K posts. We call this dataset the *Instagram 95K dataset*.

Finally, from the *Instagram 95K dataset*, we filtered out those posts with Instagram-4sq matched venues, obtaining 55,342 posts with at least one hashtag in our F&D dictionary. We call this dataset the *Instagram 55K dataset*.

In summary, our data is rich in terms of covered period (5 years) and detailed associations between individual posts and venues where they were created. Table 1 summarizes the filtered F&D datasets used in the rest of the paper. Depending on the specific analysis, we will use the corresponding dataset.

4. FOOD & DRINK ITEM DICTIONARY DESIGN

4.1 Data-Driven Dictionary Creation

Hashtags describe photos and their context, and in the case of eating-related posts they are often used to name food elements in pictures. For instance, a photo posted in Gruyere, Switzerland can have as caption: *"Feel #happy in #Gruyere. Have lunch with #cheese, #rosti at #fancy restaurant with #friends"*. In this example, #rosti and #cheese are food hashtags, while references to the location, its social context, and the user's mood are also provided.

We defined a data-driven dictionary of food and drink items as follows. As mentioned in the previous section, we started with the 65K dataset, which contains posts at 4sq food venues, so their hashtags potentially contain names of food and drink items. From this dataset, we extracted the 2,500 most frequent hashtags. Second, we defined a coding system with five hashtag categories: 0 (non-food-or-drink items, such as #geneva, #picoftheday, etc.); 1 (definite food items, such as #fondue, #cheese, etc.); 2 (definite drink items, such as #espresso, #cappucino, etc.); 3 (food-related items, such as #dinner, #lunch, etc.); 4 (drink-related items, such as #coffeeholic, #drunk, etc.). In a third step, the first author manually labeled all 2,500 hashtags according to this coding system.

Table 2 shows that only 338 (13.5%) of the top 2,500 hashtags from photos taken at 4sq food venues indeed correspond to F&D items. Furthermore, an additional 353 hashtags (14.1%) correspond to F&D-related concepts. Those F&D-related hashtags play an important role as semantic indicators of F&D events, e.g. breakfast, lunch, or healthyeating. In other words, they represent self-reported labels that indicate specific eating events. The remaining 72.4% of hashtags are about other topics. The manual coding process shows that several of these

Category	Non-food	Food	Drink	Food-related	Drink-related
# of hashtags	1805	255	83	297	56
Total	1805	338		353	
Percent	72.4%	13.5%		14.1%	

Table 2. Result of manual coding of top 2,500 hashtags at 3,245 food venues in the 65K dataset.

extra hashtags correspond to venues names, feelings of the users, current locations, etc.

The 338 F&D hashtags reflect common social media trends like frequent grammar variations (e.g. singular vs. plural), and the multilingual nature of Switzerland, a country with four official languages (German, French, Italian, Romansh) and English as lingua franca [32]. Hence, we group the original 338 hashtags into 184 F&D items. The 184 items correspond to 142 food items (F), 20 non-alcohol items (NA), and 22 alcohol items (A). The 20 non-alcohol items include beverages such as coffee, tea, and soft drinks, while the 22 alcohol items include beer, wine, cocktail, and so on. The list can be seen in Table 4. To assess differences with previous work, we compare our data-driven F&D dictionary with the ones that are publicly available [5, 26]. Interestingly, we found that they match only in 30.0% and 73.7% of elements, respectively, which highlights the interest to study world regions other than the US, which have certain globalized trends but also their own culinary variations.

4.2 Visual Validation

Previous work has shown that social media images at F&D venues cover more than just food [33]. For our dataset, we perform a validation of a sample of the F&D item dictionary to understand how much the corresponding images indeed depict such food items. First, 30 of the 184 items are randomly chosen. Second, we randomly sampled 50 pictures for each of these 30 items. Third, we defined a three-value coding system to indicate F&D item-to-image correspondence: *true* if items definitely correspond to the image content; *false* if items do not correspond at all to any image content; and *unclear* if there is an apparent connection between item and image content but we cannot be sure about it, e.g., sugar is likely part of a cake but we do not see sugar explicitly. The first author manually labeled all 1500 images (30 items x 50 pictures) according to this coding system. The results show 1,066 true cases (71.1%), 256 false cases (17.1%) and 178 unclear ones (11.8%). Unsurprisingly, food items normally used as ingredients in prepared dishes (e.g. sugar, pistachio, and mango) have the lowest visual correspondence, as sometimes they can be hardly recognizable as a separate food item. On the other hand, some items have higher visual correspondence, such as tiramisu, sashimi, or tart. In summary, this validation step highlights that our F&D item dictionary, while clearly useful, has a built-in level of uncertainty due to the way in which hashtags are created in Instagram, compared to custom-made methods to collect food labels and eating events, where people are specifically asked to label what they eat [12, 38].

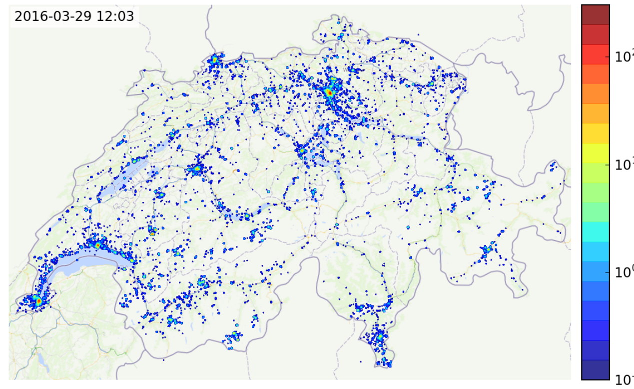

Figure 1. Spatial distribution of the Instagram 95K F&D dataset.

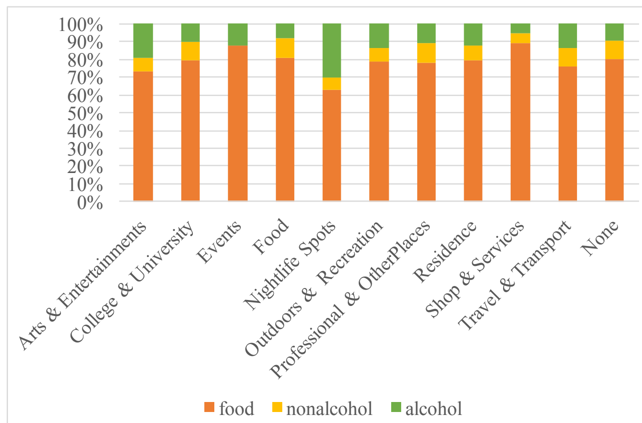

Figure 2. Percentage of food (F), non-alcohol (NA), and alcohol (A) at ten 4sq venue categories in the Instagram 55K dataset.

5. FOOD & DRINK PATTERN ANALYSIS (RQ1)

In this section, we examine the 184 F&D items on the Instagram 95K dataset. In some parts, we will mention if we use the Instagram 55K dataset.

5.1 Spatio-Temporal and Demographic Patterns

5.1.1 Spatial Patterns

Figure 1 shows the spatial distribution of F&D posts in Switzerland. Unsurprisingly, most posts come from the largest cities (Zurich, Geneva, Basel, Lausanne, Bern) and across the various linguistic regions.

We are also interested in the mean distribution of F&D posts per individual at the ten 4sq venues categories. This accounts for the bias due to frequent contributors. We examine the Instagram 55K dataset. Figure 2 shows the corresponding percentages at each venue category. Alcohol is on average most often reported at nightlife venues, which are bars, pubs, and clubs. In 8 categories, alcohol is reported with a percentage above 10% of all posts at such venues. However, the absolute number of posts at each venue category is not evenly distributed. The top 4 categories are food (21,005), travel & transport (9,687), outdoor & recreation (9,575) and nightlife spots (6,026).

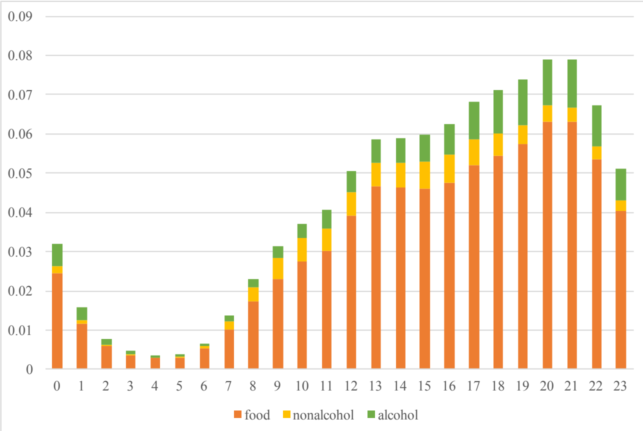

Figure 3. Hourly distribution of F&D items in the 95K dataset.

5.1.2 Temporal Patterns

Based on the GMT time of posts, we estimate the posts' timestamp in Swiss local time. Then, we calculate the hourly distribution of food, alcohol, and nonalcohol for each user. Next, we obtain the mean for all users. Figure 3 shows that the distribution of food consumption increases until the evening along with alcohol, and both of them decrease late at night. In the meanwhile, non-alcohol keeps stable during daytime from morning to afternoon. Generally, F&D pictures posted by Instagram users peak around 20:00 - 21:00 and have a local peak around 13:00 - 14:00. This means that the number of posts around dinner time are the highest followed by other eating occasions. Our results are somewhat similar to those reported in [35] using 4sq checkins in terms of main peaks, although an exact comparison is not possible. Other papers have not reported temporal patterns [5, 26, 31].

5.1.3 Demographic Patterns

Among the 42K users in the Instagram 95K dataset, there are 36K available user profile links, by examining the public availability of link "https://www.instagram.com /username/" where "username" is the self-declared user name. Many users had changed their usernames or closed their account when this was implemented. We also checked the public availability of a hyperlink to the user profile picture. Then, we use the Face++ API[3] to infer attributes of users (gender, age, and ethnicity) by linking public user profile links to the Face++ server. As a result, we get 15,504 links containing one face, 732 with two faces, 41 with three faces, and 4 with four faces. We focus on the 15K results containing one face.

From the 15K user information, we calculate the distribution of user count and post count in each range of age and gender. There are 56% females (8,682) and 44% males (6,822). In detail, female users post 18K F&D pictures (52.6% of posts), higher than the 16K (47.4%) male posts. However, on average the rate of posting F&D for males is 2.4 posts/user, while for females is 2.1 posts/user.

Regarding the age of users, the percentage of users with 35-49 year-old estimated age is the highest (50%), then 18-34 (28%) and 50-64 (19%). This result is interesting because it does not match the general demographics of American Instagram users as reported in respected surveys [16], in which young users (59%) almost double 30-49 year-old users (33%). This difference could be partly due to errors in the Face++ age estimation, but also to other factors including our focus on users who post about F&D. This question is a subject of future work.

[3] https://www.faceplusplus.com/attributes/

5.2 Eating Event Patterns

Literature in nutrition science has investigated how people define meals [21]. It is known that people label eating events employing situational factors (e.g. where and when eating takes place) [37, 25], and that how people label their meals affects what they actually eat [29]. In our case, meals are defined by the users themselves through the use of hashtags. We investigate eating events in this section.

5.2.1 Daily Meal Analysis

Hourly distribution of daily meals. We turn our attention towards daily meals: breakfast, lunch, dinner, brunch, as well as tea time and coffee time, which are often talked about (i.e., self-reported) on Instagram. From the pool of 353 F&D-related hashtags discussed in Section 4 (table 2), several of them refer to daily meals explicitly. The first author manually categorized these hashtags. They are shown in Table 3, and correspond to 11,168 posts by 6,125 users. If we only take into account the 3 main meals (9,298 posts), breakfast corresponds to 22.7% of posts, lunch to 30.2%, and dinner to 46.9%. There is uncertainty in a few hashtags (e.g. #diner can correspond to a type of restaurant or a misspelling of dinner). Breakfast posts are 7.5% (absolute) less frequent than lunch posts, and 24.9% (absolute) below dinner posts. As a relative point of comparison, the menuCH study [8] involving over 2,000 individuals, found through surveys that 5.2% of the population never have breakfast, followed by 2.2% who never have lunch, and 0.6% who never have dinner. Regarding the temporal patterns, Figure 4.a shows the hourly distribution of the six meals, with peaks for breakfast between 09:00-10:00, lunch around 13:00, brunch between 12:00-13:00, and dinner around 21:00. Tea time peaks around 16:00, and coffee time has two peaks in the mid-afternoon and morning. The hourly patterns for meals are intuitive based on observation of everyday life in Switzerland, although leaning towards the later side of what one could expect. However, note that hashtags related to daily meals can be used outside their expected time (e.g. breakfast can be mentioned in the late afternoon). Users post in this way for a variety of reasons, e.g. lack of internet connection, which makes them post at a later time.

Distribution of daily meals over the week. Figure 4.b shows the distribution of self-reported meals over the week. As patterns, breakfast and brunch have an increase on weekends (highest on Sunday). For dinner, Saturday is the day with most posts; while for lunch the most popular day is Friday.

Co-occurrence of F&D Items in Daily Meals. Based on the co-occurrence (within the same Instagram posts) of the 184 F&D dictionary items and the six daily meals, we plot wordclouds to reveal popular items used in each meal in Figure 5. People use coffee, egg, fruits, and croissant for breakfast,

Meal	F&D Related Hashtags	Post
Breakfast	breakfast, petitdejeuner, frühstück	2,119
Lunch	lunchtime, lunch, lunchbreak, lunchwithaview, lunchdate, businesslunch, pranzo	2,811
Dinner	dinnertime, diner, dinner, finedining, birthdaydinner, dining, dinnerfortwo, helvtidiner, americandiner, abendessen, ужин	4,368
Brunch	sundaybrunch, brunch, brunchtime	924
Coffee Time	coffeetime, coffetime, coffeebreak	570
Tea Time	teatime, afternoontea	376

Table 3. Daily meals: defining hashtags and frequency.

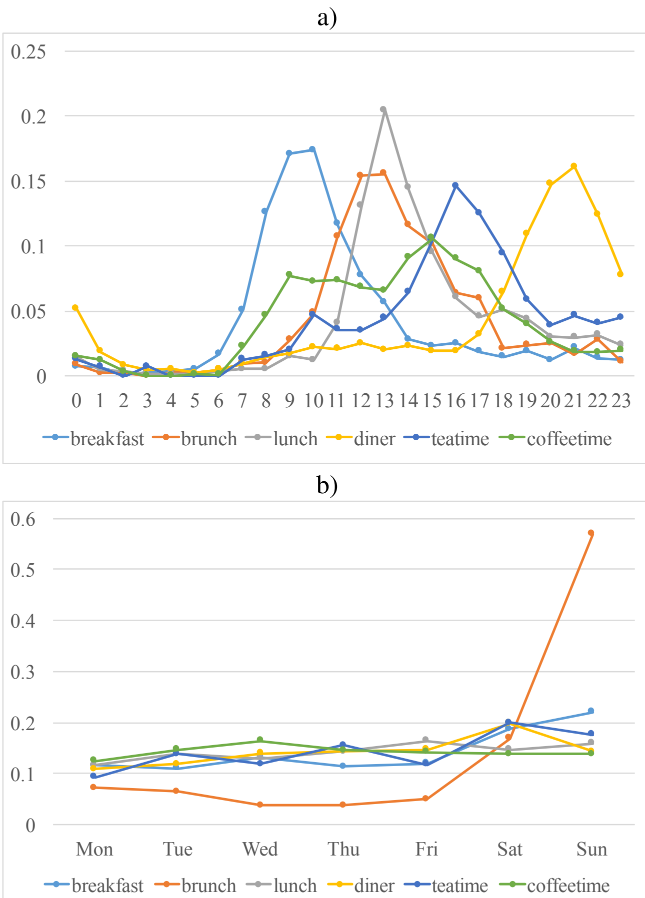

Figure 4. Temporal distribution of daily meals over (a) 24 hours; (b) days of the week.

while coffee, latte, crepes, and meat for brunch. Lunch has salad, vegetables, coffee, pasta, and burger, while dinner has wine, beef, cheese (including the fondue Swiss traditional dish), and dessert. Coffee time and tea time are characterized by the corresponding beverages and cake. Note that while the use of specific F&D items serves as illustration, we will rely on food categories in the next section as a more parsimonious description of consumed food and drinks.

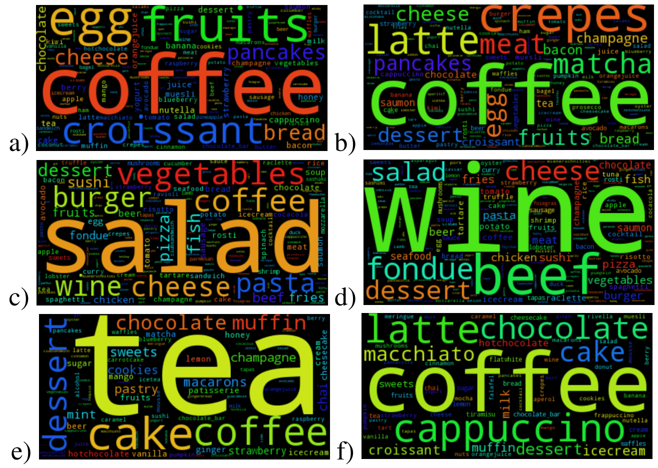

Figure 5. Wordcloud of F&D dictionary items in a) Breakfast. b) Brunch. c) Lunch. d) Dinner. e) Tea Time. f) Coffee Time.

Figure 6. Wordcloud for a) F&D dictionary items for Healty food-related hashtags. b) F&D dictionary items for Unhealthy food-related hashtags.

5.2.2 Healthy And Unhealthy Food Post Analysis

As discussed in Section 2, Instagram users comment on their own eating practices, labeling their posts as healthy (or otherwise) and using the app to keep their health goals [11]. We now investigate this phenomenon in our dataset. From the manual coding results in Table 2, we use hashtags related to both healthy food and unhealthy food to collect posts that use them. Posts are marked as healthy food if they contain at least one of the following hashtags: healthy, healthyfood, goodfood, eatclean, cleaneating, healthyeating, eathealthy, healthychoices, eatwell, fitfood, and gesund. Posts are marked as unhealthy food if they contain at least one of these hashtags: junkfood, burgerlove, burgerporn, instaburger, fastfood, mcdonalds, burgerking, and holycow (the last three being fast food chains). As a result, we obtain 3,450 healthy food posts and 666 unhealthy food posts. Figure 6 shows the occurrences of our F&D dictionary within these posts. Healthy food relates to vegetables, salad, and fruits, while (unsurprisingly) unhealthy food relates to burger, fries, and cocacola. Recent work on Instagram has shown connections between Instagram hashtags and public statistics of obesity in the US [26]. Our results suggest that there could be space to study this kind of connection for the Swiss case. As a first step, in Section 7, we investigate an approach to automate the classification of these types of posts.

6. FOOD & DRINK FFSVO CATEGORY ANALYSIS (RQ1)

6.1 Mapping F&D Items to Categories

The 184 F&D item dictionary (FDI) was manually mapped into the 19 F&D categories defined by the Swiss Federal Food

Safety and Veterinary Office (FFSVO) [4] (Federal Department of Home Affairs) by the first author, using both local knowledge and web search (e.g. wikipedia). Table 4 shows the distribution of Instagram F&D items over the FFSVO categories. Prepared dishes, sweets, vegetables, and milk & dairy products are the top food categories. Alcohol and non-alcohol drinks are also well represented. In contrast, the special foods FFSVO category (that includes food for gym or stimulants) does not exist in the dictionary. In the rest of this section, we focus on the other 18 F&D categories.

6.2 Food & Drink Category Patterns

6.2.1 Overall Distribution

We are interested in how the FFSVO categories are represented on Instagram, and how this compares with existing government data, provided among others by the Swiss Federal Statistical Office (FSO). Similarly to the previous charts, we first calculate the distribution over FFSVO categories for each user in the Insagram 95K dataset. Then, we obtain the mean distribution over all users. Note that these numbers will differ from those shown in Table 4 as they are generated from raw post counts. As a separate data source, the Swiss FSO [1] provides data on food consumption in kilograms of raw products per head per year. We group food into FFSVO categories, add the consumed amount for each category (in kilograms) and then estimate the distribution over all categories. Table 5 shows the distributions for Instagram and government data. Note that an exact comparison is not possible, given the different sources of data (hashtags counts in one case, kilograms of consumed food in the other). However, they allow for a comparison of general trends. Unsurprisingly, Instagram is biased towards certain categories, including alcohol, sweets, and prepared dishes. In contrast, certain categories are not directly represented in the official stats of FSO like prepared dishes and non-alcoholic drinks.

We see that the top 5 categories in government data are milk & diary products, fruit, vegetables, alcohol, and cereal & potatoes & starch. On the other hand, the top 5 categories on Instagram are alcohol, sweets, milk & dairy products, non-alcohol, and prepared food. It seems that what people consume in everyday life is different than what people share on Instagram. The top 5 FFSVO categories on Instagram data account for 80.0% of the probability mass, while for the official stats the top 5 categories account for 81.9%. In daily life, people consume 16.8%, 13.3%, 11.9% of cereal products, fruit, and vegetables, respectively. At the same time, Instagram users post 1.5%, 4.4%, 4.5% for the same three categories. This trend echoes media reports that state that what people post does not accurately reflect what people actually consume. There is a performative aspect to this practice, where users post what they want themselves or others to see [11]. In the Swiss case, it seems to be often sweets (19.2%) and alcohol (23.8%). The relative overabundance of these categories could partly explain why previous studies looking at connections between Instagram posts and health problems like obesity at county levels in the US have been successful at finding significant correlations [26], even though Instagram does not appear to accurately reflect true consumption patterns.

Figure 7. Bhattacharyya distributional distance between 18 F&D categories and four daily meals. The lower the distance value (violet), the more the category is used in a specific meal.

6.2.2 Main Meals and FFSVO categories

We use the Bhattacharyya distributional distance to compute the distance between the hourly distribution of each of the main meals (breakfast, brunch, lunch, and dinner) and the hourly distribution of each of the FFSVO categories (see Figure 7). For each meal type and FFSVO category, the lower the distance value, the more similar the temporal pattern is. Based on this computation, breakfast has close distance to bread, flakes, & breakfast cereal, eggs, fruit, and does not have close distance to the rest of the categories. In contrast, dinner is close to almost all categories except non-alcohol and bread, flakes, & breakfast cereal. Lunch stays somewhat in the middle with respect to many food categories. Brunch has closer distance to fruit, eggs, and non-alcohol.

6.2.3 Gender and FFSVO categories

Using the automatically inferred gender from user profile photos described in Section 5.1.3, we compute the relative posting rate between males and females for each of the FFSVO categories (see Figure 8). This number indicates whether a given F&D category is most popular among men or women (for each category, the two rate values add up to one). The relative rate of male posting is higher than female posting for meat substitutes, meat & offal, alcohol, and sausages & cold meats. The rate is similar for males and females w.r.t. non-alcohol. For the rest of categories, the relative female posting rate is higher. The plot shows a differentiated trend of posting of certain categories based on gender that could be worth investigating in more depth in future work.

As a point of comparison in terms of general trends, we can compute a similar plot for certain food categories using the data in kilograms per head per year from the menuCH study [2, 8] (not shown for space reasons). In this case, except for non-alcohol, males have a higher relative consumption than females for all other categories. Once again, an exact comparison is not possible given the different measures used in each case.

Category	% FDI	Samples of food drink items (FDI)
Bread, Flakes & Breakfast Cereals	2.2	bagel, bread, croissant, muesli
Cereal Products, Pulses & Potatoes	6.0	vermicelles, noodles, pasta, penne, spaghetti, carbonara, ramen, gnocchi, ravioli, rice, risotto
Egg	1.0	egg, meringue
Fats & Oils	0.5	butter
Fish Crustaceans	4.9	saumon, tuna, shrimp, lobster, oyster, mussels, caviar, fish, seafood
Fruit	7.1	fruits, avocado, pineapple, strawberry, raspberry, apple, mango, berry, kiwi, blueberry, coconut, banana, passionfruit
Meat & Offal	3.8	beef, ribs, lamb, pork, chicken, duck, meat
Meat Substitute	0.5	tofu
Milk & Dairy Products	8.2	yogurt, cheese, parmesan, mozzarella, raclette, fondue, cream, latte, milk, chai, milkshake, cappuccino, macchiato, frappuccino, flatwhite
Nuts, Seeds & Oleaginous Fruit	1.1	nuts, pistachio
Prepared Dishes	16.3	sandwich, clubsandwich, soup, salad, wienerschnitzel, pho, padthai, paella, pizza, sashimi, sushi, maki, crepes, tapas, carpaccio, nachos, fajitas, guacamole, tartare, curry, escargot, flammkuchen, couscous, antipasti, kebab, falafel, dessert, matcha, piadina, burger
Sausages & Cold Meats	2.7	ham, sausage, bacon, salami, foiegras
Savoury Snacks	1.1	fries, chips
Sweets	13.0	carrotcake, tiramisu, macarons, cake, pancakes, cheesecake, muffin, donut, patisserie, tart, pastry, cookies, waffles, gingerbread, honey, sugar, sweets, chocolate, nutella, caramel, chocolate bars (Lindt, Toblerone, etc.), icecream, mousse, pannacotta
Various	1.6	ketchup, sauce, wasabi
Vegetables	11.4	mushrooms, asparagus, tomato, pumpkin, vegetables, onion, cucumber, spinach, truffle, rucola, mint, edamame, lemon, lime, vanilla, ginger, olives, cinnamon, pesto, potato, rosti
Alcohol Drink	12.0	alcohol, wine, prosecco, beer, tequila, lillet, grappa, aperol, spritz, cocktail, margarita, gintonic, gin, bellini, mojito, champagne,cognac, whisky, liquor, aperitif, vodka, sake
Non-alcohol Drink	6.5	coffee, tea, bubbletea, mocha, hotchocolate, cocacola, rivella, lemonade, gazosa, icetea, juice, orangejuice

Table 4. Distribution of 184 F&D dictionary items (FDI) (non-normalized over users) over the FFSVO food categories for the Instagram 95K dataset.

FFSVO Category	Instagram (%)	FSO (%)
Bread, Flakes & Breakfast Cereals	0.76	N/A
Cereal Products, Pulses & Potatoes	1.59	16.85
Egg	0.56	1.28
Fats & Oils	0.08	2.91
Fish Crustaceans	2.42	0.93
Fruit	4.48	13.36
Meat & Offal	3.43	5.81
Meat Substitute	0.04	N/A
Milk & Dairy Products	13.47	29.05
Nuts, Seeds & Oleaginous Fruit	0.13	1.05
Prepared Dishes	10.54	N/A
Sausages & Cold Meats	1.23	N/A
Savoury Snacks	0.55	N/A
Special food or Stimulants	N/A	1.45
Sweets	19.20	4.65
Various	0.14	N/A
Vegetables	4.51	11.97
Alcohol Drink	23.85	10.69
Non-alcohol Drink	12.99	N/A

Table 5. Distribution of F&D FFSVO categories in the Instagram 95K dataset and official statistics (FSO). For some categories, data is not available (N/A).

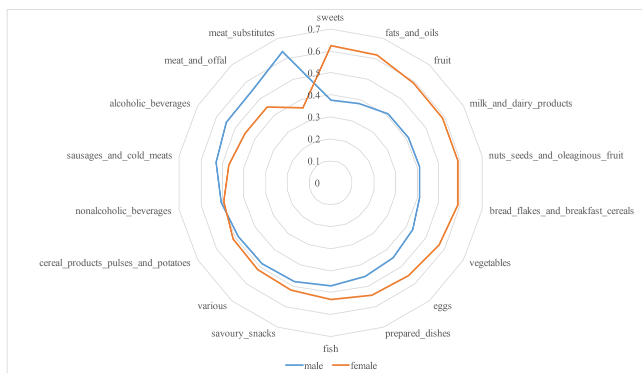

Figure 8. Distribution of male and female mentioning F&D categories.

7. CLASSIFICATION OF EATING EVENTS (RQ2)

The previous sections showed how Instagram users employ hashtags to mention eating events (e.g. lunch) and what they think about them (e.g. unhealthy). This form of self-report is interesting for two reasons. First, users make use of their own internal definitions to choose the hashtags they attach to their posts; in other words, they decide on their own what they call breakfast or healthy. Second, this bottom-up practice results in labeling eating events in a useful way for supervised learning. In this section, we follow such approach for two eating event classification tasks: a six-class daily meal classifier (breakfast, lunch, brunch, dinner, tea time, coffee time), and a binary healthy vs. unhealthy food classifier. In both cases, we study *content* features (184 F&D items and 17 food categories), *context* features (time of day, day of week, and 4sq venue category), and *social* features (likes and comments), with the goal of understanding their individual and combined discriminative power.

7.1 Classification Method

Random Forest (RF) is a well-known supervised learning method for classification [23]. It builds up multiple decision trees, and the output of classification is the mode of the results over all individual trees. RF is able to deal with numerical data and categorical data (typically handled by using factors or one-hot encoding). In the reported experiments, we use one-hot encoding, and set parameters as ntree = 500 and mtry as recommended by [23]. We use repeated 10-fold cross validation over 5 times for accuracy evaluation, i.e., 9 data folds are used for training and 1 data fold is used for testing. This procedure is repeated 5 times.

7.2 Feature Extraction

Features are extracted from textual data of Instagram posts and 4sq venues. We group them into six groups: F&D items (F), F&D categories (FC), context (C), social (S), picture caption statistics (P), and Foursquare statistics (4sq). Note that we treat the picture caption statistics separate from the actual hashtag content (F), and the venue category (part of context C) separate from the specific venue statistics, so as to have a cleaner representation of content and context. We summarize all features used for classification in Table 6.

7.3 Classification Results and Discussion

For the two classifications tasks, the datasets are imbalanced. In case of daily meals, discussed in Section 5.2.1, there are 6 classes spreading from 4,368 dinner posts to 376 teatime posts. For experiments, we decide to keep the original number of posts for the six classes. In the case of healthy and unhealthy posts, discussed in Section 5.2.2, we decide to balance the dataset. We randomly chose 666 healthy posts from the 3,450 available posts such that healthy posts and unhealthy posts are equally represented.

Healthy vs. Unhealthy Classification. In term of individual features, the F&D item feature (F) is the best feature, with 83.2% accuracy. Then, F&D category (FC) is the second best individual feature with 79.3% accuracy. This result is expected as specific food items are related to the corresponding posts being labeled as healthy or not (recall Fig. 6). The combination of F and FC decreases slightly to 82.6%. In term of feature group combinations, the combination of content and context (F+FC+C+P) provides the highest accuracy with 85.8%. This suggests that time and venue category provide additional discriminative power. The rest of the features do not contribute to further improve classification performance.

Six Daily Meal Classification. We have 11,168 daily meal posts with the following distribution: breakfast (19.0%), brunch (8.3%), dinner (39.1%), lunch (25.2%), coffee time (5.1%), and tea time (3.4%). A majority class baseline (labeling everything as dinner) thus represents an accuracy of 39.1%. In terms of individual feature groups, context (C) is the best feature with 60.7% accuracy. In principle, time is intuitively a good cue to discriminate among some daily meals, although the problem is not trivial given the overlap in time, place, and menu items that many of these meals can have. Furthermore, the F&D item feature (F) is the second best feature with 56.6% accuracy. In term of combinations, a group that integrates content and context (F+C+P) provides the highest accuracy with 61.7%. It is interesting to see that the use of food items can indeed complement the context information albeit slightly. The results also show that the classification task remains open for future performance improvements.

In terms of the most relevant features from the RF for each of the two tasks, the top 10 sub-features for healthy vs. unhealthy are: burger (F&D item), the number of food related tags, prepared dishes (F&D category), salad (F&D item), the number of hashtags, savoury snack (F&D category), pizza (F&D item), food venue category, fries (F&D item), and unknown (venue category). In contrast, the top 10 sub-features for classifying daily meals are: time of the day, tea (F&D item), coffee (F&D item), cappuccino (F&D item), pancakes (F&D item), wine (F&D item), croissant (F&D item), day of the week, fruits (F&D item), and eggs (F&D item). For both classification tasks, the lists of most relevant features seem meaningful.

In summary, we have shown that F&D content and context are indeed informative features for eating event classification in the two tasks we studied. In this sense, by studying the social media setting, our work adds to recent work in ubiquitous computing that is examining how to automatically identify eating events from mobile and wearable sensors [36, 28].

8. FINAL DISCUSSION AND CONCLUSION

In this paper, we set out to study Instagram food and drink posting in a particular national context. We close the paper by summarizing the answers we found to the two research questions we posed, and by discussing limitations and future directions.

Our first question (RQ1) inquired about the types of food and drink consumption patterns that could be mined from Instagram data generated in Switzerland, and about how these patterns compared to national statistics. We have shown that, starting from a large and longitudinal dataset of Instagram posts and the definition of a data-driven F&D item dictionary, several patterns related to spatial distribution, temporal distributions, basic demographics, food categories, and eating events can be extracted. We found that the F&D item dictionary, around which the whole study was conducted, plays a key role, and is not identical to dictionaries created in other western countries in previous work. This highlights the importance of understanding the national context under which social media studies on food and drink are conducted. We will publish the list of F&D items and categories we collected in our dataset. We also found that broad comparisons with national statistics on the subject are possible but not exact. Despite this limitation, some of the biases of Instagram data appear evident through these comparisons. This points out towards caution when investigating social media data as a proxy for everyday life. At the same time, this does not remove the value of understanding food and drink consumption on Instagram as a specific social media practice.

Feature	Description	Type	Group Feature
hour	Time of the day (in minutes) when the picture is posted	numeric	Context (C)
day	week days when the picture is posted	numeric	Context (C)
vennuecat	4sq venue category where the picture is posted	categorical (10)	Context (C)
likes	number of likes of the picture	numeric	Social (S)
comments	number of comments of the picture	numeric	Social (S)
userInPhotos	number of userInPhotos in the picture	numeric	Social (S)
filter	filter user uses for the picture	categorical (44)	Social (S)
tags	number of tags in the picture	numeric	Picture Caption (P)
captions	number of words of captions in the picture	numeric	Picture Caption (P)
foodtags	number of F&D hashtags mentioned in the picture	numeric	Picture Caption (P)
foodrelatedtags	number of F&D related hashtags mentioned in the picture	numeric	Picture Caption (P)
checkinsCount	number of checkins of 4sq users at venue	numeric	Foursquare (4sq)
usersCount	number of users did check-ins at venue	numeric	Foursquare (4sq)
tipCount	number of tips posted by 4sq users at venue	numeric	Foursquare (4sq)
F	184 F&D binary vector	categorical (184)	F&D items (F)
FC	18 F&D category vector	categorical (18)	F&D Categories (FC)

Table 6. Features for classification of eating events.

Feature	Acc(%)
Baseline	50.0
F	83.2
FC	79.3
C	71.1
4sq	62.8
P	61.7
S	58.0
F + FC	82.6
F + FC + P	84.7
F + FC + P + C	**85.8**
F + FC + P + C + 4sq	84.9
F + FC + P + C + 4sq + S	85.3

Table 7. Classification results for healthy and unhealthy (N= 1332).

Feature	Acc(%)
Baseline on majority class	39.1
F	56.6
FC	54.7
C	60.7
4sq	43.5
P	34.9
S	39.1
C + F	61.6
C + F + P	**61.7**
C + F + P + 4sq	61.3

Table 8. Classification results for six daily meals (N= 11,168).

Our second question (RQ2) inquired whether content and context features could be used to automatically classify eating events. We have shown that a number of features could be defined from the insights obtained from the descriptive analysis, and a random forest approach was able to classify healthy vs. unhealthy posts with 85.8% accuracy, and could also classify six daily eating occasions with 61.7% accuracy, both with a combination of content and context features.

Future work will investigate a few issues that we could not fully studied here. Two of them have to do with demographics. Our analysis showed that slightly older people accounted for the majority of users who posted F&D in our Swiss dataset. Our analysis also showed that there is a gender difference in the posting rate about specific food categories (like alcohol and sweets). Understanding these issues in detail could involve a mixed-method approach, where data analytics would be complemented by qualitative approaches. A third issue has to do with is the role of image content (via automatic recognition) in the refinement and extension of some of our current analysis, e.g. to understand the social context under which food and drink are posted and talked about.

ACKNOWLEDGMENTS

This work has been funded by the EPFL Integrative Food Science and Nutrition Center through the Social Food and People project, the Swiss National Science Foundation through the Dusk2Dawn Sinergia project, and a Swiss Government Excellence Scholarship. We thank Darshan Santani for discussions.

REFERENCES

1. 2016. Food and Agriculture Pocket Statistic 2016. (2016). https://www.bfs.admin.ch/bfs/en/home/statistics/agriculture-forestry/food.html Accessed: 2017-08-21.
2. 2017a. menuCH: Resultats concernant la consommation alimentaire. Consommation des differents groupes d'aliments. (2017). https://www.blv.admin.ch/blv/fr/home/lebensmittel-und-ernaehrung/ernaehrung/menuch/menu-ch-ergebnisse-ernaehrung.html Accessed: 2017-08-23.
3. 2017b. Results on Food Consumption Survey. (2017). https://www.blv.admin.ch/blv/fr/home/lebensmittel-und-ernaehrung/ernaehrung/menuch.html Accessed: 2017-08-07.
4. 2017c. Swiss Food Composition Database. (2017). http://naehrwertdaten.ch/request?query=TopCategoryList&xml=MessageData&xml=MetaData&xsl=ListCategories&lan=de&range=0-19 Accessed: 2017-08-01.
5. Sofiane Abbar, Yelena Mejova, and Ingmar Weber. 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*. ACM, 3197–3206.
6. Morteza Akbari Fard, Hamed Hadadi, and Alireza Tavakoli Targhi. 2016. Fruits and Vegetables Calorie Counter Using Convolutional Neural Networks. In *Proceedings of the 6th International Conference on Digital Health Conference*. ACM, 121–122.
7. Christina Berg, Georgios Lappas, Alicja Wolk, Elisabeth Strandhagen, Kjell Torén, Annika Rosengren, Dag Thelle, and Lauren Lissner. 2009. Eating patterns and portion size associated with obesity in a Swedish population. *Appetite* 52, 1 (2009), 21–26.
8. Murielle Bochud, Angéline Chatelan, Juan-Manuel Blanco, and Sigrid Beer-Borst. 2017. Anthropometric characteristics and indicators of eating and physical activity behaviors in the Swiss adult population. In *Results from menuCH 2014-2015 (2017)*.
9. Géraldine M Camilleri, Caroline Méjean, France Bellisle, Valentina A Andreeva, Valérie Sautron, Serge Hercberg, and Sandrine Péneau. 2015. Cross-cultural validity of the Intuitive Eating Scale-2. Psychometric evaluation in a sample of the general French population. *Appetite* 84 (2015), 34–42.
10. Jingjing Chen and Chong-Wah Ngo. 2016. Deep-based ingredient recognition for cooking recipe retrieval. In *Proceedings of the 2016 ACM on Multimedia Conference*. ACM, 32–41.
11. Chia-Fang Chung, Elena Agapie, Jessica Schroeder, Sonali Mishra, James Fogarty, and Sean A Munson. 2017. When Personal Tracking Becomes Social: Examining the Use of Instagram for Healthy Eating. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*. ACM, 1674–1687.
12. Felicia Cordeiro, Elizabeth Bales, Erin Cherry, and James Fogarty. 2015. Rethinking the Mobile Food Journal: Exploring Opportunities for Lightweight Photo-Based Capture. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems (CHI '15)*. ACM, New York, NY, USA, 3207–3216.
13. Aron Culotta. 2014. Estimating county health statistics with twitter. In *Proceedings of the 32nd annual ACM conference on Human factors in computing systems*. ACM, 1335–1344.
14. Kiyah J Duffey and Barry M Popkin. 2011. Energy density, portion size, and eating occasions: contributions to increased energy intake in the United States, 1977–2006. *PLoS medicine* 8, 6 (2011), e1001050.
15. Daniel Fried, Mihai Surdeanu, Stephen Kobourov, Melanie Hingle, and Dane Bell. 2014. Analyzing the language of food on social media. In *Big Data (Big Data), 2014 IEEE International Conference on*. IEEE, 778–783.
16. Shannon Greenwood, Andrew Perrin, and Maeve Duggan. 2016. Social media update 2016. *Pew Research Center* 11 (2016).
17. Christopher Holmberg, John E Chaplin, Thomas Hillman, and Christina Berg. 2016. Adolescents' presentation of food in social media: An explorative study. *Appetite* 99 (2016), 121–129.
18. Yoshiyuki Kawano and Keiji Yanai. 2014. Food image recognition with deep convolutional features. In *Proceedings of the 2014 ACM International Joint Conference on Pervasive and Ubiquitous Computing: Adjunct Publication*. ACM, 589–593.
19. Daniel Kershaw, Matthew Rowe, and Patrick Stacey. 2014. Towards tracking and analysing regional alcohol consumption patterns in the UK through the use of social media. In *Proceedings of the 2014 ACM conference on Web science*. ACM, 220–228.
20. Tomasz Kusmierczyk and Kjetil Nørvåg. 2016. Online Food Recipe Title Semantics: Combining Nutrient Facts and Topics. In *Proceedings of the 25th ACM International on Conference on Information and Knowledge Management*. ACM, 2013–2016.
21. Rebecca M Leech, Anthony Worsley, Anna Timperio, and Sarah A McNaughton. 2015. Understanding meal patterns: definitions, methodology and impact on nutrient intake and diet quality. *Nutrition research reviews* 28, 1 (2015), 1–21.
22. Linna Li, Michael F. Goodchild, and Bo Xu. 2013. Spatial, temporal, and socioeconomic patterns in the use of Twitter and Flickr. *Cartography and Geographic Information Science* 40, 2 (2013), 61–77.
23. Andy Liaw, Matthew Wiener, and others. 2002. Classification and regression by randomForest. *R news* 2, 3 (2002), 18–22.

24. Janne Lindqvist, Justin Cranshaw, Jason Wiese, Jason Hong, and John Zimmerman. 2011. I'M the Mayor of My House: Examining Why People Use Foursquare - a Social-driven Location Sharing Application. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '11)*. ACM, New York, NY, USA, 2409–2418.
25. David Marshall and Rick Bell. 2003. Meal construction: exploring the relationship between eating occasion and location. *Food quality and Preference* 14, 1 (2003), 53–64.
26. Yelena Mejova, Hamed Haddadi, Anastasios Noulas, and Ingmar Weber. 2015. # FoodPorn: Obesity patterns in culinary interactions. In *Proceedings of the 5th International Conference on Digital Health 2015*. ACM, 51–58.
27. Austin Meyers, Nick Johnston, Vivek Rathod, Anoop Korattikara, Alex Gorban, Nathan Silberman, Sergio Guadarrama, George Papandreou, Jonathan Huang, and Kevin P Murphy. 2015. Im2Calories: towards an automated mobile vision food diary. In *Proceedings of the IEEE International Conference on Computer Vision*. 1233–1241.
28. Mark Mirtchouk, Christopher Merck, and Samantha Kleinberg. 2016. Automated estimation of food type and amount consumed from body-worn audio and motion sensors. In *Proceedings of the 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing*. ACM, 451–462.
29. Patricia Pliner and Dragana Zec. 2007. Meal schemas during a preload decrease subsequent eating. *Appetite* 48, 3 (2007), 278–288.
30. Francesco Ragusa, Valeria Tomaselli, Antonino Furnari, Sebastiano Battiato, and Giovanni M Farinella. 2016. Food vs Non-Food Classification. In *Proceedings of the 2nd International Workshop on Multimedia Assisted Dietary Management*. ACM, 77–81.
31. Jaclyn Rich, Hamed Haddadi, and Timothy M. Hospedales. 2016. Towards Bottom-Up Analysis of Social Food. In *Proceedings of the 6th International Conference on Digital Health Conference (DH '16)*. ACM, New York, NY, USA, 111–120.
32. Darshan Santani and Daniel Gatica-Perez. 2013. Speaking Swiss: languages and venues in Foursquare. In *Proceedings of the 21st ACM international conference on Multimedia*. ACM, 501–504.
33. Darshan Santani and Daniel Gatica-Perez. 2015. Loud and trendy: Crowdsourcing impressions of social ambiance in popular indoor urban places. In *Proceedings of the 23rd ACM international conference on Multimedia*. ACM, 211–220.
34. Sanket S Sharma and Munmun De Choudhury. 2015. Measuring and characterizing nutritional information of food and ingestion content in instagram. In *Proceedings of the 24th International Conference on World Wide Web*. ACM, 115–116.
35. Thiago H Silva, Pedro OS Vaz de Melo, Jussara M Almeida, Mirco Musolesi, and Antonio AF Loureiro. 2014. You Are What You Eat (and Drink): Identifying Cultural Boundaries by Analyzing Food and Drink Habits in Foursquare.. In *ICWSM*.
36. Edison Thomaz, Irfan Essa, and Gregory D Abowd. 2015. A practical approach for recognizing eating moments with wrist-mounted inertial sensing. In *Proceedings of the 2015 ACM International Joint Conference on Pervasive and Ubiquitous Computing*. ACM, 1029–1040.
37. Brian Wansink, Collin R Payne, and Mitsuru Shimizu. 2010. "Is this a meal or snack?" Situational cues that drive perceptions. *Appetite* 54, 1 (2010), 214–216.
38. Lydia Zepeda and David Deal. 2008. Think before you eat: photographic food diaries as intervention tools to change dietary decision making and attitudes. *International Journal of Consumer Studies* 32, 6 (2008), 692–698.