Abstract

Food insecurity in the UK has captured public attention. However, estimates of its prevalence are deeply contentious. The lack of precision on the volume of emergency food assistance is made even more ambiguous due to peer-to-peer food sharing systems (e.g. OLIO). While these initiatives exist as a solution to food waste not food poverty, they are nonetheless carrying a hidden share of the food insecurity burden, with the socio-economic status of technology-assisted food sharing donors, volunteers, and recipients remaining obscure. In this article we examine the relationship between food sharing and deprivation generally, then use machine learning to develop a predictive model of food insecurity based on aggregated food sharing behaviours by OLIO users in the UK. We demonstrate that data from food sharing systems can help quantify a previously hidden aspect of deprivation and we make the case for a reformed approach to modelling food insecurity.

Keywords:

Food Sharing, Food Waste, Food Insecurity, Economic Deprivation, Machine Learning, Sharing Economy

Identifying Food Insecurity in Food Sharing Networks via Machine Learning

1. Introduction

Food insecurity is a remarkably persistent problem in the United Kingdom. However, existing estimates of its prevalence are deeply contentious. Despite being the fifth wealthiest country in the world (Credit Suisse, 2019), inequality endures to such an extent that many people cannot afford basic provisioning. Recent research (Taylor-Robinson et al., 2013; Ashton, Middleton and Lang, 2014; Loopstra, Reeves, Taylor-Robinson et al., 2015; Dowler and Lambie-Mumford, 2015b; Dowler and Lambie-Mumford, 2015a) suggests an increasing number of people are facing food hardship, many experiencing 'in-work poverty' where salaries fail to cover even basic expenditure.

Though the UK government measures an *Index of Multiple Deprivation*[1], which includes income, employment, and health, there exists a dearth of reliable large-scale data on the number of people unable to reliably procure enough food for themselves and their families (Department for Communities and Local Government, 2015). As a result, politicians routinely question the veracity of available data, resulting in political inaction. Emergency food assistance statistics provided by foodbanks have been criticised for having an ulterior political agenda or not actually reflecting food insecurity (e.g. Walker, 2017).

The lack of precision on the total volume of emergency food assistance is made even more ambiguous due to food-surplus redistribution organisations (e.g. Fareshare) and peer-to-peer food sharing systems (e.g. OLIO, see Michelini, Principato and Iasevoli, 2018). These initiatives exist as a solution to food waste not food poverty, but they are nonetheless carrying a hidden share of the food insecurity burden (Baron et al., 2018). More importantly, food waste and food poverty are 'ethical problems' at the core of two sustainable development goals. Their alleviation, therefore, is a 'morally relevant aspect of modern life', as well as 'a global joint

[1] The IMD is a publicly available composite measure of deprivation calculated using data on income, employment, health, education, housing and crime rate aggregated at catchment-area level, with 42,619 such neighbourhoods in the UK, with different nations formalising 'neighbourhoods' in slightly different ways

political effort' for all actors, from farmers, food retailers, consumers, to public welfare and third sector voluntary system (Galli, Cavicchi and Brunori, 2019, p.1).

The food industry, like many others, is being transformed by peer-to-peer technology. Sharing and access-based consumption have dramatically grown in popularity, and consequently new business models continue to emerge which deviate from traditional linear supply chains (e.g., Botsman and Rogers, 2010; Breitsohl, Kunz and Dowell, 2015; Bucher, Fieseler and Lutz, 2016; Harvey et al., 2019; Hellwig et al., 2015; Lamberton and Rose, 2012; Scaraboto, 2015; Schor and Fitzmaurice, 2015). Yet despite the rapid growth of food surplus sharing, the socio-economic status of technology-assisted food sharing donors, volunteers, and recipients is obscure (Harvey et al., 2019). This represents a glaring hole in the social scientific understanding of food insecurity generally. Indeed, as previous work suggests, 'these demographics are obscure and there are no reliable statistics available' on the number of people in food insecurity who may be turning to redistribution and sharing services (Harvey et al., 2019).

This article is organised as follows: we examine the relationship between food sharing and deprivation generally, before subsequently using machine learning to develop a predictive model of food insecurity based on aggregated food sharing behaviours exhibited by OLIO users in the UK. In the following section, we review extant literature on food insecurity and food surplus sharing in the UK. We discuss some of the current problems facing managers and policymakers trying to understand the scale of the problem of food insecurity. The third section provides an overview of the research approach taken. We outline the rationale for using OLIO data and machine learning to predict food insecurity and assist those most in need. The fourth section provides an exposition of the results and discussion. In the final section we conclude the paper by highlighting the theoretical contributions of the approach and make the case for a reformed approach to modelling national food insecurity statistics.

We demonstrate that data generated through food sharing initiatives can help document a previously hidden and un-quantified aspect of economic deprivation. Beyond describing this aspect of deprivation, the second key contribution of this article is the development of the first estimate of population-level risk of food insecurity for all neighbourhoods in the UK. Though the observed data originates from a single platform, and is necessarily biased to those with the ability to access it, its extensive coverage means that it produces a picture of food insecurity at a more fine-grained level than previously obtained, and at a greater geographical extent that can be achieved from traditional surveys. Furthermore, as the underlying model can be automated across contexts, the method is much more cost and time effective than surveys of food insecurity in the UK, which have only been conducted sporadically.

2. Theoretical background and research context

2.1 The intersection of food surplus, food sharing, and food insecurity

'*Food bank use is up almost fourfold since 2012, and there are now about 2,000 food banks in the UK, up from just 29 at the height of the financial crisis. Not only does the government not measure food insecurity, but a Minister dismissed the significance of foodbank use as being only occasional and noted that foodbanks exist in many other western countries. The clear implication was that their rapid growth in the UK should not be seen as cause for concern, let alone for government action*.' (United Nations Special Rapporteur, 2019, p.17)

Though the UK maintains an Index of Multiple Deprivation (IMD), including economic factors such as income and unemployment, the constituent variables do not focus specifically on access or consumption of food. This is problematic because accurate and non-partisan measurement of the problem is a prerequisite to any informed political intervention. Furthermore, research on food security (Loopstra and Tarasuk, 2015) has already drawn attention to the limitations of existing empirical data sources (e.g. foodbank donations).

Food insecurity has serious long-term health impacts on the people it affects. Consequences include inadequate nutrient intake, micronutrient deficiencies, malnutrition, diet-related chronic disease, and exacerbated mental health issues (Alaimo, Olson and Frongillo, 2001; Whitaker, Philips and Orzol, 2006; Hackett, Melgar-Quinonez and Alvarez, 2009). Food insecurity is especially pernicious for the cognitive development and health of children who grow up in households where food is scarce. These increasingly prevalent issues have prompted health experts in the UK to suggest*, 'This has all the signs of a public health emergency that could go unrecognised until it is too late to take preventive acti*on' (Taylor-Robinson et al., 2013). Though the Government has recently announced plans to begin small-scale measurement of household food insecurity in 2021 as part of existing household research, these statistics will not reveal the depth or prevalence of the issue at large. Consequently, there is an urgent need to create new measures, which can quantify food insecurity longitudinally. This is particularly necessary in geographic regions where food insecurity is made worse by 'food deserts' i.e. locations with poor access to affordable food. See for instance, the recent work completed on food deserts in the UK by the The Social Market Foundation (2018).

Food insecurity and food waste are generally conceptualised, studied, and managed separately, but a growing number of academics and practitioners have called for recognition of their interrelation and coexistence, particularly in the global North (FAO, IFAD, UNICEF, WFP and WHO, 2017; Baron et al., 2018; Galli, Cavicchi and Brunori, 2019). Every year millions of people experience hunger, despite the fact that 1.3 billion tonnes of edible food surpluses are disposed as waste (Gustavsson, Cederberg and Sonesson, 2011). In the UK, during 2015, 7.3 million tonnes (worth approximately £13 billion) of edible food was thrown away by households (House of Commons Environment, Food and Rural Affairs Committee, 2017). These figures suggest that resolving the dilemma of scarcity within abundance requires a 'global joint political effort', with United Nations states committed to meeting these two sustainable development goals by 2030:

- SDG2: 'Ending hunger and ensuring access by all people, the poor and the vulnerable to safe, nutritious and sufficient food',
- SDG12: 'Ensuring sustainable consumption and production patterns', with the goal of 'halving per capita food waste and reducing food losses' (United Nations Development Program, 2016).

At the same time, their alleviation is a 'morally relevant aspect of modern life' for all actors, from farmers to food retailers. Consumers cannot escape blame either, with new marketing approaches needed to influence individual behaviour (Galli, Cavicchi and Brunori, 2019, p.1). The intersecting evils of waste and insecurity have consequently led many organisations to launch food waste initiatives, many of which aim to reconfigure food-surplus supply chains through new forms of redistribution and sharing.

Food sharing is an ancient and fundamental behaviour enacted in all human cultures (Enloe, 2003; Jaeggi and Gurven, 2013). However, like food insecurity, very limited quantitative data has historically been available to describe the practice of food sharing between non-kin. This absence of data is beginning to change, as demonstrated by Harvey et al. (2019), whose study of OLIO's peer-to-peer network shows digital data offers a starting point to capture and analyse food-sharing statistics at scale. Initial research suggests there tends to be an imbalance in the roles performed by food sharing donors and recipients, as many people using OLIO primarily act as either a donor or a recipient, regardless of whether they share their own food or on behalf of local companies. The implication is that there is likely to be a subset of recipients who are turning to food redistribution and sharing services like OLIO due to hunger, rather than just the environmental motivation of reducing food waste through sharing. Indeed, as Harvey et al. (2019) argue: *'As more food surplus becomes managed through applications there is an opportunity for policymakers to work with organisations to improve . . . the identification of vulnerable people experiencing food hardship. If OLIO, like other systems that enable donations or other forms of altruistic sharing, can calculate a dependency index (i.e. what proportion of the network are dependent on other members of the network) this could give a profound insight into how macroeconomic policy is affecting the food consumption habits of consumers.'*

2.2 The challenge of measuring food insecurity in the UK

Household food insecurity has been understood as a symptom of the 'interacting pathologies of household poverty, community disadvantage, and the actions of the food industry', and diagnosed using 'landscape metaphors of food deserts, food swamps and food brownfields' (Thompson, D. Smith and Cummins, 2019, p.2). Despite offering a valuable theoretical framing, these conceptualisations of food environments have not yet been reflected in routine measurements of food insecurity in the UK. Instead, national statistics have commonly been used to estimate the prevalence of food insecurity in the UK.

The IMD, the most commonly used indicator of social and material deprivation in the UK, may identify areas at risk of food insecurity through some of its domains reflecting neighbourhood levels of unemployment, income and health deprivation (D. Smith, Thompson et al., 2018). Another important body of work has focused on the relationship between the spatial accessibility to healthy food sources, diet quality and neighbourhood deprivation (Cummins, D. Smith, Taylor et al., 2009; D. Smith, Cummins et al., 2010; Black et al., 2014; Clary et al., 2010).

Food insecurity surveys, often deployed in other countries, are not routinely used across all four countries in the UK. Four exceptions stand out. In 2004, the Food Standards Agency (FSA) commissioned the Low Income Diet and Nutrition Survey, covering the 15% most deprived households in the UK (Nelson et al., 2007). From 2016, the FSA started reporting levels of food insecurity in England, Wales and Northern Ireland in the bi-annual Food and You Survey (NatCen Social Research, 2017). Most recently, the Mayor of London commissioned a survey on the severity of food insecurity in Greater London (London Datastore, 2019). These surveys have captured the experience of destitution- from running out of food to going without eating for days, using a validated tool widely used in high-income countries- the US Department of Agriculture Adult Food Security module (United States Department of Agriculture Economic Research Service, 2017).

Based on a sample of 3,318 adults, the 2016 survey showed that one in five adults in the UK experienced food insecurity, with 10% reporting living in moderate or severe destitution. In an examination of the magnitude and severity of food insecurity captured in the 2004 and 2016 surveys, Loopstra, Reeves and Tarasuk (2019) found that this phenomenon was most pronounced for people with low-incomes, whose prevalence rose from 27.7% in 2004, to 45.8% in 2016. Among factors associated with food insecurity, Loopstra, Reeves and Tarasuk (2019) documented younger age, ethnicity, low education, disability, unemployment, and low income. Similar levels of destitution were observed in the 2019 survey of 6,601 Londoners, with 21%, or 1.5 million adults estimated to be in moderate or severe food security (London Datastore, 2019). A significant development for research and policy was made in 2019, when the Department for Work and Pensions included food insecurity questions in the annual Family Resources Survey. This survey addresses important limitations of previous approaches- it is large scale covering 20,000 households across all four countries in the UK (Loopstra, 2020). Its limitation, however, is that it will be at least 2021 before these results will be available for analysis. The resulting lack of data not only hinders progress on public policy in the UK, but renders existing local statistics susceptible to questioning (Walker, 2017).

In brief, existing measures of food insecurity in the UK often rely on small localised samples, or national statistics and global regression models, which reflect old data sources and assume that relationships are invariant across time and spatially distant communities. It would be beneficial, therefore, to complement these composite measures of deprivation with quarterly updated information on benefit claiming at different levels of aggregation (Department for Work and Pensions, 2017). More so, given the proliferation of food aid networks (Loopstra, Reeves, Taylor-Robinson et al., 2015), the location and prevalence of foodbanks could be used as a proxy for levels and distribution of food insecurity. Research by Loopstra, Reeves, Taylor-Robinson et al. (2015) has shown that 'Trussell Trust food banks are more likely to open in local authorities characterised by cuts to central welfare, higher rates of unemployment and higher rates of benefit sanctions' (D. Smith, Thompson et al., 2018, p.22). Moreover, research into the relationship between food banks and food insecurity in Canada has shown that it is individuals in severe destitution who seek help from an extended networks of support, including family, friends and community agencies (Tarasuk, St-Germain and Loopstra, 2019). This comes with the acknowledgement that, due to their third sector set-up, the daily running of foodbanks is often 'based on community resources and local social networks - not an objective measure of need or population characteristics' (D. Smith, Thompson et al., 2018, p.22). Moreover, despite their undeniably vital role in alleviating food insecurity, they are limited in improving the diets of low-income people, due to their provision of long-life, pre-processed food (Thompson, D. Smith and Cummins, 2019).

In order to fully understand current dietary practices in the context of increasing food insecurity and reliance on networks of food aid, a more holistic approach to characterising food environments is needed. Conceptualising food security as 'access to enough food for an active, healthy life' (McEntee, 2009, p.355) would help understand food environments as including both top down and participatory approaches to supply, as well as integrating notions of spatial, economic and informational access.

Access to longitudinal proprietary data opens possibilities for understanding real-time and fine-grained food consumption behaviours, and linking them to socio-demographic and attitudinal data at different levels of aggregation (Strong, 2015). Supermarket loyalty card records and behavioural data collected from food-sharing platforms like OLIO only offer a fragmented view of food consumption and dietary practices. These records, however, are amenable to food basket methodology (Anderson et al., 2007), or more advanced behavioural and semantic segmentation like topic modelling (Hruschka, 2014) that would complement our understanding of current dietary practices.

Finally, while national statistics offer insights on geographic propinquity and socio-economic affinities, social network analysis can reveal further insights about the homophily and interdependencies arising in local networks of food aid (Ducruet and Beauguitte, 2014; McPherson, Smith-Lovin and Cook, 2001). Specifically, by modelling the activities of individuals acting within a food network it is possible to investigate connections between behaviour and often hidden social challenges such as deprivation and food poverty. To this end, we explore not only socio-demographics, but the behaviours and inter-dependencies revealed in food sharing app usage. This addition yields three distinct dimensions to examine: 1. users' neighbourhood characteristics; 2. users' behavioural repertoires; and 3. users' position within a food sharing network topology as a whole. In the following section, we introduce the background to OLIO, whose data makes such analysis possible.

2.3 OLIO – A solution to food waste through sharing

OLIO was founded in 2015 by Tessa Clarke and Sasha Celestial-One. The company provides a service that 'connects neighbours with each other and with local businesses so surplus food can be shared, not thrown away. This could be food nearing its sell-by date in local stores, spare home-grown vegetables, bread from your baker, or the groceries in your fridge when you go away' (OLIO, 2019). OLIO is available for free through Apple and Android application stores and is also accessible via web browsers. The service is the most popular of its kind in the world. At the time of writing (June 2020) there are 2 million registered users who have shared 5,344,356 portions of food across 53 countries. OLIO is especially popular in the UK where the organisation was initially created.

OLIO's popularity is due in large part to a network of volunteers who collect food surpluses from local businesses and then redistribute them to people who request specific items. The organisation thus fosters C2C relations, B2C relations, and B2V2C relations (Business to Volunteer to Consumer). There are a wide variety of stakeholders within the OLIO ecosystem, each of whom are interested in reducing food waste, but each also stand to gain a variety of other benefits e.g. increased social interaction between neighbours, the opportunity to gain free fresh food for consumers, and reduced disposal costs for businesses.

The interface provides similar functionality to most popular social networking sites (e.g.: scrolling feeds, profile pages, direct messaging, as depicted in Figure 1) but focuses on enabling people to share food as easily as possible (as shown in figure below). The application also enables people to request food from their local neighbourhood rather than responding directly to items shared with the network. The request functionality within OLIO provides a window into the motivations people have for acquiring food surpluses.

Research on OLIO user motivation has not yet received serious academic study. OLIO has been discussed briefly in previous work (Lazell, Magrizos and Carrigan, 2018; Carrigan, 2017; Schanes, Dobernig and Gozet, 2018; Schanes and Stagl, 2019), and the network structure has been examined in-depth by Harvey et al. (2019), but no existing empirical studies have examined the prevalence of food insecurity within the network. Indeed, to our knowledge the relation between food sharing and food insecurity has never been studied extensively or at scale, due to a lack of longitudinal data.

However, OLIO is keen to understand how widespread food insecurity is across the UK, and where possible improve the design of their service to assist those most in need. To be clear though, OLIO was designed as a solution to fight food waste, not poverty. And there are a growing number of researchers (Caplan, 2017; Caraher and Furey, 2017) who have rightly criticised the idea that food surplus sharing solutions should have to carry any of the burden of food insecurity. Food insecurity is a consequence of political choices. There is a real danger that emergency food assistance and food redistribution services become normalised despite being a 'band-aid to more deep-rooted problems of poverty' (Caraher and Furey, 2017). But at present, though these arguments about normalisation are convincing, they remain at least partly conjectural due to the lack of available statistical evidence examining the relation between food surplus redistribution and food insecurity generally. In the following section we outline a method to begin to redress this issue.

— [Figure 1 here] —

3. Research approach

- *Research Question 1: What is the relation between deprivation and food-sharing behaviour?* There is yet to be any comprehensive study of the relation between deprivation and food sharing between non-kin. This research question is a direct response to work by Caplan (2017) and Caraher and Furey (2017) who have called for greater scrutiny of the relation between food surplus sharing/redistribution and deprivation. The answer has managerial implications for the way peer-to-peer sharing organisations understand, manage, and promote their relations with the broader infrastructure of emergency food assistance. But it also has theoretical relevance for the way in which food insecurity is conceptualised generally.
- *Research Question 2: Can food insecurity be predicted from food-sharing behaviour?* This research question is a response to the work of Harvey et al. (2019), to ask whether it is possible to predict instances of food insecurity among OLIO users, and thus understand the wider prevalence of the issue. The marketing management implications of this question relate to the moral imperative of data custodians to be cognisant of vulnerable consumers.

The overall aim of the research is to form a predictive model, and hence shed light on, the prevalence of food sharing users who are likely to be experiencing food insecurity. Specifically, the approach focuses on distinct forms of observed and inferred data which can be used to build a predictive model, these include:

- Network topology: who interacts with whom - i.e. how OLIO users choose each other to share food;
- Neighbourhood characteristics: secondary socio-economic markers of the geographic area in which users are located;
- User behavioural repertoires: the ways that people use the OLIO application in order to perform a variety of tasks i.e. temporally varied and qualitatively distinct forms of human-computer interaction.

There is emerging evidence that data driven decision making can be beneficial for firm performance (Sivarajah et al., 2017; Kubina, Varmus and Kubinova, 2015; Brynjolfsson, Hitt and H. H. Kim, 2011). Collating huge amounts of data ostensibly opens a portal to better knowledge, performance and change prediction. As George, Haas and Pentland (2014, p.323) observe, 'The fine-grained nature of big data offers opportunities to identify these sources of change...'. Machine learning is driving this transformation through predictive and descriptive analytics (e.g. more sophisticated segmentation and summarisation). However, computational business research is still not common; in marketing it is rare, despite some siren voices, e.g. Lilien and Rangaswamy, 2004, and some isolated specific applications, e.g. Boone and Roehm, 2002. In fact, the data-driven discovery paradigm is not salient in marketing at all; an area historically dominated by hypothetico- deductive research or interpretive methods, Ehrenberg (1988) being a long standing exception that proves the rule. This is a missed opportunity since in practice marketing is increasingly analytics driven, increasingly inductive, and predictive models are useful for theory generation. Marketing academia should not be immune from these developments. As other fields of study have acknowledged, the growing volume of transactional data creates the possibility of new methods for computational social science (Savage and Burrows, 2007; Lazer et al., 2009).

The detection of vulnerable consumers is ethically essential for OLIO. Such end-users are potentially dependent on this re-distribution network. Data custodians have a duty to consider those at risk of consequences from the decisions and actions informed via analytics. If they can be identified, then they should. Data should not just be used for commercial expediency; it should also be deployed in the interests of consumer welfare. Indeed, organisations, such as OLIO, that subscribe to this ethos are likely to enjoy much greater levels of trust and operate more sustainable models of data usage.

3.1 Data collection

In order to provide a nuanced understanding of individuals' experience of food insecurity, distinct forms of observed and inferred data were used for assembling the three dimensions characterising OLIO's users, as detailed in the following section and summarised in Table 1.

— [Table 1 here] —

The network topology and behavioural repertoires views were assembled from the dataset obtained from OLIO. This dataset can be best summarised as conversations among users, whereby some offered items, others requested them, with food exchanged offline being marked as successfully 'picked up'. These conversations also include a category of 'wanted items' - immediate and specific requests. The records were collated by querying transaction data logged on OLIO's servers, with each showing anonymised user identifiers, a timestamp indicating when the account was created or modified, and registration latitude and longitude.

The transactional data contains a three-year period from 9th July 2015 until 30th October 2018. There are 141,129 unique items listed in the dataset, of which 102,239 were requested 238,622 times and 99,604 were exchanged offline by 41,811 users of the network in Britain. A similar pattern was found in the roles adopted by individuals using the OLIO platform, with most using it to request or donate food and only a small percentage engaging in both roles, as detailed in Table 2. An emerging role on the platform is played by volunteers who collect donations from local affiliated stores and distribute them into the community, despite accounting for about two percent of OLIO's active user base. More importantly, only half of the individuals requesting food articles on the platform were at the receiving end of donations, with much demand going unmet, or alternative sharing practices developing offline. In the context of reducing food poverty, the sharing and request practices enabled by OLIO merit more research. To understand the social configurations enabling solutions to this issue and to reflect behavioural repertoires on the platform, this study focuses on a sample of 26,980 users who have requested, or requested and donated food.

— [Table 2 here] —

The spatial distribution of OLIO users is illustrated in Figure 2. Previous work in the UK has examined how spatial accessibility to healthy food influences the relation between neighbourhood deprivation and diet quality. Such work often uses the UK official measure for material hardship, the IMD. This study also leveraged this metric, estimating the level of hardship of each OLIO user via the IMD score of the neighbourhood they registered in. We recognise the nosiness of such an assignment, while believing it provides the best such estimate currently available without turning to costly and potentially invasive surveying procedures. Deprivation estimates for users were normalised as a score between 1 and 100. Overall, and as shown in Figure 3, OLIO users tend to reside in areas of higher than average hardship - across all the dimensions captured by the IMD, apart from Educational Deprivation, and with Living Environment deprivation being particularly pronounced.

— [Figure 2 here] —

— [Figure 3 here] —

In addition to the IMD, this study utilised a second indicator of deprivation based on users' probability of soliciting benefits. For this we again used data inferred from neighbourhood catchment areas, specifically Employment and Support Allowance (ESA) and Pension Credit (PC) scores (from May to August 2018), both of which are income-related benefits (Department for Work and Pensions, 2017).

Finally, information on infrastructure, including the location of food stores, bus stops and foodbanks, as well as each user's distance to these locales, was collated from national census and open data scraping, as detailed in Figure 4. The location of 12,009 food stores was scraped from the Food Standards Agency's platform. This platform provides data on food hygiene ratings or inspection results for businesses including restaurants, pubs, cafés, takeaways, hotels and other places where consumers eat, as well as supermarkets and other food stores (Food Standards Agency, 2019), while the location of 2,212 foodbanks and centres was mapped based on data from the Trussell Trust network (The Trussell Trust, 2018), as well as from the Independent Food Aid Network (Independent Food Aid Network, 2018). Information on bus stop locations and each user's distance to the nearest one was calculated based on national public transport access node (NaPTAN) schemas and guidance from the UK government, adding up to 444,462 stops (GOV.UK, 2018).

— [Figure 4 here] —

3.2 Research methods

The research approach involved three steps: first, three user dimensions were assembled; these were then used as basis for exploration of the association between deprivation and food-sharing using correlation analysis, with machine learning being leveraged to model and identify instances of food insecurity.

3.2.1 Assembling the three dimensions of food insecurity

For the network topology dimension, exploratory social network analysis in NetworkX – Python language library (Hagberg, Swart and S. Chult, 2008) - was used and focused on measures of degree distribution, centrality and clustering, indicative of interdependency and homophily.

The second dimension brought together the distinct behavioural repertoires users expressed on the platform through descriptive statistics and behavioural segmentation. Commonly used in consumer research, segmentation involves dividing 'a heterogeneous market into relatively homogeneous segments' (Foedermayr and Diamantopoulos, 2008, p.223) to gain a deeper understanding of customer preferences, needs and wants. Approaches include cross-tabulation, cluster analysis, non-negative matrix factorization (NMF) and latent Dirichlet allocation (LDA), and have previously been applied to shopping records, text corpora (DiMaggio, Nag and Blei, 2013) and location data (Eagle and Pentland, 2009) to represent a phenomenon of interest as a linear combination of components. In the case of OLIO, we used NMF to identify meaningful temporal patterns in regard to users’ soliciting, as well LDA to summarise diet preferences from users’ requests, resulting in each user being summarised as a mixture of temporal and food soliciting routines.

Finally, we assembled the neighbourhood characteristics view by associating measures of deprivation and information on infrastructure to each user’s account based on their neighbourhood of registration.

3.2.2 Exploring associations between deprivation and food-sharing behaviour. Modelling and predicting food insecurity

These three dimensions formed the basis for exploration of the association between deprivation and food-sharing behaviour using correlation analysis on a sample of 26,980 users who have requested or requested and donated food. Finally, a sub-sample of 421 users was prepared for training and testing four classification algorithms, in order to model food insecurity based on the three collated user dimensions and identify instances of food insecurity in the broader sample.

Using food solicitation messages, we identified instances of legitimisation, beneficiary focus and emotional appeal as a basis for classifying soliciting users as being in food hardship or not. These strategies have commonly been researched in the context of charitable-giving and found to play a significant role in donations. For example, Cialdini and Schroeder (1976) investigated the effect of the legitimisation of minimal donations as strategy to eliciting donations, while research into beneficiary focus has shown that its effect on charitable behaviour depends on the salience of the appeal (White and Peloza, 2009). Moore and Harris (1996), too, has documented the use of dramatic emotional appeals, such as fear and guilt, for grabbing the attention of potential donors. In classifying users as being in hardship or not, this protocol, too, follows commonly use surveying approaches and relies on users’ self-declared food insecurity, rather than our own perception of manipulative intent in regards to their solicitations.

We found semantic nuance in this sample’s solicitation messages: some specifically asked for a product - perhaps after having run out, or when trying a new diet. Others sought left-over food, long shelf life products, even nearing ‘use by’ date. Most commonly, however, these users' solicitation strategies were telling of their level of food hardship. Legitimising phrases such as 'Any food greatly appreciated, not eaten for 2 days and have another week before benefits cut in' seems to carry minimal cost for compliance, while 'making it difficult for people to reject the request' (Shearman and Yoo, 2007, p.273). On a platform where the salience of social norms is high, highlighting how a donation would benefit both donor and recipient is likely to be even more effective: 'Reduce food waste and help feed my family'. These latter messages also exemplify emotional appeals, with family (ageing parents and children) often cited as beneficiaries of these requests.

Other users tried to distance themselves from the act of soliciting food, as indicated by use of passive voice and no first-person pronouns (e.g., 'Any food for a few days meals wanted'). Self-declared food insecurity, as well as the intensity of these strategies in users' soliciting messages have led to the identification of 222 users in food hardship and 199 as not being in hardship. This comes with the acknowledgement that the classification may underestimate the prevalence of users in food insecurity, as some may be tentative to broadcast their hardship to the network. However, given that the predictive model is based on multiple behavioural features, those users who might feel uncomfortable requesting food may nonetheless exhibit similar behavioural repertoires (e.g. an emphasis on requesting food), and would thus be recognised by the model.

Machine learning methods have been applied to recognise patterns and classify records in a wide variety of applications. More importantly, they allow the leveraging of commercial and public data for achieving business or social good goals (Engelmann, Goulding and G. Smith, 2018). To investigate the hierarchy of factors predictive of food insecurity, a classification task was formulated based on the sub-sample of 421 users. Competing models were generated based on the three complementary dimensions, and the performance of each model in predicting food insecurity status was tested via cross-validation. Though commonly used for classification tasks, Support Vector Machines (SVM), Random Forests (RF), Adaptive Boosting (AdaBoost) and k-Nearest-Neighbour (kNN) give qualitatively different results. A support vector machine transforms the feature space into a higher dimension, 'where it finds a hyperplane that separates the data by class' (Han, Pei and Kamber, 2011, p.393).

Alternatively, Random Forests and AdaBoost are better suited at handling non-linear relationships. They are also examples of ensemble methods, which combine a series of learned models to create an improved composite classification model, with Random Forests being comparable in accuracy to AdaBoost, yet showing robustness to error and outliers. Finally, the k-Nearest-Neighbour method provides an intuitive approach to classification. It relies on learning by analogy, comparing a given test tuple (a new user's attributes) with training examples that are similar to it by using a distance metric, such as Euclidean distance (Han, Pei and Kamber, 2011).

The prediction task underpinning these models was formulated as a binary classifier (self declared food insecurity or not). The performance of each model was assessed using five-fold cross-validation in conjunction with the training set to determine the optimal hyper-parameters then measuring performance on the held-out test set. Finally, an extensive variable selection analysis was performed via principal component analysis and ANOVA F-value feature ranking.

The motivation behind this was two-fold. Firstly, principal component analysis (PCA), Latent Dirichlet Allocation (LDA) and non-negative matrix factorization (NMF) are dimensionality reduction approaches commonly applied to databases of facial images (Lee and Seung, 1999), corpora of documents (Gautam and Shrestha, 2010) and behavioural data (Eagle and Pentland, 2009). Not only do they manage abundant variables in a dataset by learning linear combinations of components, they also enhance 'the performance of the pattern analysis algorithms' (Chaki and Dey, 2019, p.11170). In this case, it was possible to maintain a good approximation of the original data by only using the strongest principal components. Secondly, variable selection had the role of automatically ranking and selecting attributes in the data according to their relevance to the prediction task. They did not alter the original representation of the variables, but merely selected a subset. In brief, this helped decrease dimensionality, avoid over-fitting, reduce training time and improve model interpretability. The following sections describe these steps in more depth.

3.2.3 Network topology dimension

In this section we provide a short description of the 45 features making up the three inter-related user dimensions input for food insecurity prediction, as detailed in Table 3. For assembling the network topology view we focused on measures of degree distribution, centrality and clustering. Measures of centrality are a relevant stream of research for identifying key players, as well as structural measures of social capital in a social network (Bonacich, 1972; Freeman, 1979; Borgatti, Jones and Everett, 1998). For the OLIO network, too, we were interested in examining how individual users were best positioned to reach all others more quickly and control the flow of donations between them. While degree centrality captures the number of nodes one is connected to, it does not account for the structure of the network. Intuitively, although a user might be connected to many others, it might not be in a position to reach others quickly to access donations. To capture this, we also focused on measuring one's reach - using closeness centrality - as well as the degree to which a node lies on the shortest path between two other nodes, controlling the flow in the network - using betweenness centrality (Borgatti, 2006).

— [Table 3 here] —

At the same time, what are the underlying factors that influence the formation of relations? Defined as the principle whereby users with similar characteristics tend to associate with each other (in terms of age, social status, or network standing), 'homophily' is often researched in the context of reciprocal relations and can reflect link, as well as status affinities between users (Hopcroft, Lou and Tang, 2011). In particular, we were interested in measuring status homophily using the PageRank algorithm to estimate each user's importance relative to network structure. Forms of link homophily were further explored in the neighbourhood characteristics dimension.

Unlike many undirected relations that have been modelled using network analysis, such as co-authoring, or word collocations (Iijima and Kamada, 2017), OLIO's food sharing platform affords multiple, directed behaviours. For this, we sought to understand how users vary across in-degree (illustrating the number of requests received by a potential donor) and out-degree (number of requests placed by a user) to get a sense of the intensity of one's food needs, as well as the 'burstiness' of their behaviour. If a person is irregularly, but frequently - in a short period of time - asking for food donations, this may indicate an unpredictable and disproportionate demand in comparison to other nodes in the network.

Alternatively, it may suggest that a user in a dense neighbourhood is more likely to develop more relations - place more requests - than one in a sparse community. Clustering coefficient, therefore, was another measure included in OLIO's network topology view, reflecting 'the fraction of pairs of a person's collaborators who have also collaborated with another one' (Al Hasan et al., 2006; Ravasz and Barab´asi, 2003).

3.2.4 Neighbourhood characteristics

With geographic propinquity 'creating contexts in which homophilous relations form', it is important to complement insights on homophily gleaned from social network analysis with insights derived from spatial networks (Ducruet and Beauguitte, 2014; McPherson, Smith- Lovin and Cook, 2001). We argue that physical distance between donors and receivers, as well as social status affinities play an important role in influencing the formation of relations in sharing networks. For the OLIO network we were interested to examine how users place requests based on the spatial distance between themselves and donors. At the same time, drawing on research evidencing 'negative associations between food poverty and health outcomes' (Thompson, D. Smith and Cummins, 2018), we were also interested in users' accessibility to traditional food chains and networks of aid. For this, we included distance to nearest bus stop, supermarket and food bank, as well as number of such locales as dimensions of this neighbourhood view. This comes with the acknowledgement that food access is not limited to the spatial dimension, with residents of deprived neighbourhoods in Scotland shown to have high levels of access to grocery and fresh produce stores (Cummins, D. Smith, Aitken et al., 2010). To further explore hardship beyond material metrics, the UK's composite measure of deprivation, its seven dimensions, as well as each user's probability of soliciting benefits, were included as part of OLIO users' neighbourhood view.

3.2.5 Behavioural repertoires

With emerging evidence that data driven decision making can be beneficial for firm performance (Kubina, Varmus and Kubinova, 2015; Brynjolfsson, Hitt and H. H. Kim, 2011), OLIO's behavioural records have the potential to illuminate the reasons why users may be soliciting donations (Hruschka, 2014). Temporal routines of platform use may give insights into the periodicity of users' needs, while the semantic dimensions of their donation or request records may be telling of one's motivations for using OLIO (Gautam and Shrestha, 2010; Hruschka, 2014; Shah, Kumar and K. H. Kim, 2014). If a user is occasionally requesting non-core foods, such as fresh rhubarb, or spare ounces of cacao, they may not be experiencing the same level of food insecurity as users asking for a wider range of foods, especially staple, even nearing 'use by' date products. Lastly, we were also interested in behaviours implicit to OLIO's network structure, such as number of likes one had placed on other users' listed items, as well as the number of likes they had received in turn. We include the recency, frequency and quantitative value of their requests, as well as maximum number of requests placed in a day. We also include number of received items, as well as number of collections from affiliated stores to reflect the many behaviours manifest on the platform.

3.3 Ethics

In adherence with OLIO's terms and conditions all data collated came from adults (over the age of 18). All network and behavioural data were anonymised at point of collection and all analysis is presented at an aggregate level to remove the possibility of data triangulation.

4. Results and discussion

The results follow the three dimensions laid out earlier - network topology, neighbourhood characteristics, and behavioural repertoires - to contextualise the relationship between deprivation and food sharing, and model food hardship. Reflecting earlier research on food sharing through OLIO (Harvey et al., 2019), exploratory network analysis showed that most individuals use the platform to donate or request food, with only 13% engaging in both roles.

Examination of network interdependency measures gives a more nuanced understanding of these relationships. The average degree of all users of the network is 15.52, with a standard deviation of 87.57. Further examination of in-degree (received donation requests) and outdegree (placed donation requests) indicates that there is a supply and demand imbalance in the network, with only half of the individuals requesting food articles on the platform - or 32% of users - being at the receiving end of donations. This might suggest that much demand is going unmet, or alternative sharing practices are developing offline (Table 2). More so, 86.6% of users had requested food no more than 10 times. This level of demand, however, was directed towards 94.55% of donors. The findings also indicate that OLIO's network exhibits a distribution in which a small number of individuals form a larger number of connections than the average user. This core of users who intensely donate and request on the platform have disproportionate in and out degree scores, indicative of qualitative distinctions and great variation in the way they perform these behaviours on the platform.

Degree measures only give a localised view of platform connections. To further explore aspects of interdependency, we now interrogate the network's level of clustering. This measure was shown to keenly abstract the concept of social distance, with large coefficients, positive degree correlations, and the emergence of a hierarchy of communities empirically describing 'the tendency of peers to establish acquaintances via a decreasing function of relative distance' (Bogu˜n´a et al., 2004, p.70). Measuring the number of transitive relationships between peers, each node's clustering coefficient was calculated for the undirected version of the OLIO network, leading to a mean coefficient of 0.12, with a standard deviation of 0.26 across the sample.

In Figure 5 we observe that the clustering coefficient is independent of the degree distribution, contrasting to many real networked systems showing a decreasing function of degree (Boguñá et al., 2004). The implications are clear, however: high-degree and low-degree nodes show qualitatively diverse behaviours. Low degree, high clustering nodes may denote small tight-knit communities where food is shared over short distances and neighbours have more connections among themselves. High degree, high clustering coefficient may indicate users with many sharers who are likely to have connections from similar communities. Reflecting Harvey et al. (2019)'s initial findings on the OLIO platform, these behaviours indicate a departure from ‘the simple traditional commodity lifecycle characterised by one-way flows from producer to consumer’ under the interplay of associated stores, volunteers and consumers. These network components have a clear correspondence to spatial accessibility, as 67% of requesters were located within 10 kilometres of potential donors. This is to be expected given the size of the network and the exchanges requiring parties to travel in order to share food with one other. 13.20% of requests, however, occur over larger distances. This latter finding marks a departure from traditional food sharing models based on kin-selection and exemplifies the non-linear configurations between volunteers collecting food parcels from associated stores, then reaching out to the broader community to distribute them.

— [Figure 5 here] —

While networked interactions of food sharing were found for people around the world, its motivational drivers cannot be limited to reciprocity, with ‘social contexts of lived lives’ also influencing these behaviours (Ng et al., 2013, p.7). Examination of neighbourhood characteristics, including spatial and economic access to food extend preliminary findings about food sharing. The map in Figure 2 presents the location of users in the UK, showing high densities in London boroughs, as well as in important urban centres. Perhaps unsurprisingly, this is reflected in increased spatial accessibility of food stores and public transport links, with 33.85% of users having between one and four supermarkets in their neighbourhood and 68.27% of users having up to 10 bus stops (Figure 4). Moreover, 84.80% of users have registered within one kilometre - the equivalent of a 15-minute walk or 3-minute drive - from the nearest store, 97.99% registering within three kilometres, the equivalent of a 20-minute walk or 4-minute drive. In terms of food aid accessibility, 94.89% of users were located within five kilometres of a food bank or centre, while the most distant registration was in the Scottish Highlands. As far as public transport accessibility is concerned, 99.61% of users were located within 1 kilometre of a bus stop.

In addition to geographic propinquity providing an insight into homophilous connections, we also turn to the broader social and economic affinities that have formed in the network and explore hardship beyond material metrics through UK’s composite measure of deprivation. The average difference between donors’ and

requesters' level of deprivation (IMD rank) is calculated, with an average of -0.21 and a standard deviation of 27.94 indicating that most individuals request from users experiencing varying levels of deprivation, greater and lower than their own.
The network and spatial dimensions of food sharing on the OLIO platform demonstrated 'qualitative distinction in forms of donors and recipients', as well as variation in the performance of these roles (Harvey et al., 2019). Daily request burstiness in particular challenges the traditional recourse to altruism evidenced in food-sharing research (Holme and Saram¨aki, 2013). As these users' food soliciting routines show, most use the platform sporadically but steadily, others - voraciously: a relatively small percentage of users - 6.20% - use the platform intensely to request food. Although it is presumptuous to generalise these findings to motivations, temporal segmentation of users' requests provides insights into the periodicity of their needs. As Figure 6 details, some request food arbitrarily, throughout the week. Others, on evenings and weekends only. Semantic segmentation of recipients' food records also gives an insight into latent needs, activities and diet preferences. Some users irregularly request from a large assortment of foods, others have more sporadic and specific requests, e.g., kombucha scoby and rhubarb, while other stock on staple and long-life products, as detailed in Figure 7.

— [Figure 6 here] —
— [Figure 7 here] —

Received and extended likes give further nuance of platform affordances that may be encouraging desired behaviours.

4.1 Exploring the association between deprivation and food-sharing behaviour

The inter-relations of the three user dimensions demonstrate that there is variance in the way users share food through OLIO. Although it is widely assumed that food insecurity is closely related to existing deprivation measures, correlation analysis of the dimensions making up these three views showed that this relationships is not straightforward. Perhaps unsurprisingly, middling negative correlation between living environment deprivation and number of food stores was observed ($r = -.23$, $p < .001$), indicating better spatial accessibility for neighbourhoods suffering physical decay. Moreover, high negative correlations were observed between one's probability of claiming benefits and income deprivation (r= -.77, p <.001), as well as employment (r= -.82, p <.001) and health deprivation (r= -.72, p <.001), attesting to the persistence and inter-relation of these issues. In terms of the relationship between foodbanks locations and deprivation, however, low levels of correlation were observed with the composite measure (r= .22, p <.001), income (r= .16, p <.001), employment (r= .13, p <.001) and health (r= .20, p <.001). This indicates that the location of foodbanks is not a reliable proxy for insecurity, but may instead reflect community resources and local social networks.

No relation was established between measures of deprivation and levels of food donation or sharing on the platform. This confirms that OLIO is predominantly used in diverse urban areas, with good spatial access to traditional food chains and public transport links. Although one's temporal sequence of roles on the platform was not captured in the analysis, received likes positively correlated with measures of network centrality and clustering, while number of extended likes showed middling positive correlations with measures of interdependency. Notably, the number of received likes positively correlated with PageRank (r= .52, p <.001), in-degree (r= .58, p <.001), betweenness centrality (r= .34, p <.001) and degree centrality (r= .52, p <.001), while extended likes positively correlated with in-degree (r= .18, p <.001) and out-degree levels (r= .52, p <.001).

4.2 Modelling food insecurity

The overall aim of the research is to model and identify instances of food sharing users who are likely to be experiencing food insecurity. Specifically, the approach focuses on distinct forms of observed and inferred data - collated under the three user dimensions - which were used to build a predictive model. Evaluation measures were used throughout this process, in support of feature selection, as well as for assessing the performance of competing classifiers (Sun, Wong and Kamel, 2009, p.696).

For each model, the data was split into a training (four-fifths) and test set (one-fifth) and the parameters for the four classifiers (e.g., number of trees for RF, nearest neighbour for kNN), as well as for each feature selection approach (number of components for PCA and number of best predictive features for feature ranking) selected via a grid search underpinned by five-fold cross-validation, as detailed in Table 4. Three-to-one and four-to-one training to test ratios are typical in machine learning applications, from genomics (Vabalas et al., 2019) to land use classification (Engelmann, Goulding and G. Smith, 2018). In this case we opted for the latter split in order to have more training data available for model fitting and avoid an increase in pessimistic bias of the model's performance estimate, as recommended by Raschka (2018). The performance of each model was then tested on the held-out test set, resulting in accuracy, precision and recall scores for each of the feature selection approaches. As some users experiencing food insecurity may be unwilling to broadcast their hardship to the network, it was important to evaluate the model's positive predictive value through precision scores, as well as through the combined f1 metric (Saito and Rehmsmeier, 2015).

— [Table 4 here] —

While PCA applied to the four classifiers showed encouraging levels of prediction accuracy, it was feature ranking and selection with Random Forests that showed meritorious prediction levels. As detailed in Table 5, the latter method showed the highest levels of precision and recall. Moreover, f1 scores for each model illustrate that these prediction levels were not the result of one class being disproportionately favoured over the other. While marginally so, f1 scores for users not experiencing hardship were lower - 0.75 compared to 0.76. This resonates with insights from initial exploration of the sample, whereby users' food security changes during their time on the platform. Some start by soliciting food for themselves, then transition to acting as volunteers for OLIO, collecting food from associated stores and distributing it in the community. At the same time, RF and AdaBoost, models handling non-linear interactions, showed increased prediction accuracy, indicating a complex structure of features contributing to one's food insecurity status. Admittedly, noise and access bias are inherent in natural experiments, more so when there is stigma associated with food soliciting- the behaviour we are observing. While ground truth is not readily available in this case, evaluation of the model's positive predictive performance gives us valuable insights into an under-researched phenomenon. Moreover, we suggest that future improvements of the model should include triangulation of insights based on surveys, to ensure and improve the ongoing efficacy of prediction.

— [Table 5 here] —

Of equal, if not greater importance than overall model accuracy, is the ranking of features contributing to these prediction levels, with dimensions of interdependency, clustering, as well as behavioural repertoires contributing to the prediction of food hardship (Figure 8). The number of likes one received on their food listings ranked as top predictive variable - which is not surprising given that there was a significant difference in the number of items donated on the platform by users self-declaring to be in hardship ($M= .71$, $SD= 3.09$) or not ($M= 18.17$, $SD= 63.96$, $t(419)= -4.06$, $p <.05$). This was confirmed by the selection of in-degree (received food solicitations) and number of listed food articles as top ranking features for the prediction of food insecurity status. In terms of behavioural repertoires, users also showed a significant difference in their food solicitation routines, with users in hardship predominantly requesting irregularly throughout the week - Soliciting routine 1: $M= .21$, $SD= .15$, $t(419)= -4.87$, $p <.05$.

— [Figure 8 here] —

In terms of neighbourhood characteristics, health and education deprivation, rather than the composite measure of deprivation, were ranked as top contributing factors to one's hardship status. This resonates with earlier research emphasising the role of informational access to food as mediating between economic deprivation and diet quality.

Finally, the number of connections one has on the platform, as well as their standing in the broader OLIO network were shown to carry weight in the prediction of one's food insecurity status. Independent t-test showed a statistically significant difference between the degree centrality and PageRank of users in hardship ($M= .0007$, $SD= .0009$ and $M= .00008$, $SD= .0002$, respectively) and user classified as not being in hardship ($M= .0003$, $SD= .0015$ and $M= .00002$, $SD= .00002$, respectively), with values of $t(419)= -3.92$ and $t(419)= -4.32$, $p <.05$.

5. Conclusion

5.1 Conceptual implications

Those most in need in the OLIO case are characterised by a particular profile of network usage. These elements make intuitive sense but crucially are evidenced here for the first time. Recipients in acute food need are not sharers on the whole; they are takers. Their attempts to engage in other functions (e.g. donation) are muted and sparse. Reciprocity is largely absent; sharing is therefore more akin to donation/receipt.

There is also a gatekeeper effect within the network. Recipients in acute need rely on the highly active community of volunteers prepared to travel to enable re-distribution. So, there is a dependency effect; the food insecure are dependent on the vitality of volunteers. This is a reflection of the kind of phenomena observed generally in communication networks (Hopcroft, Lou and Tang, 2011) and marketing communications in the analogue era. Gatekeepers are crucial; in the OLIO case their connectedness and relatively high centrality scores are also a reflection of donor behaviour, not just communication. Indeed, their elevated levels of communication centrality are a reflection of their importance in the redistribution network.

5.2 Methodological contribution

The IMD contains collinearities across the seven variables measured. Though it is widely assumed that food insecurity is closely related to existing deprivation measures (e.g. low income households, unemployment, benefit claiming), the results show that this relationship is not straightforward. The geographic distribution of food insecurity is heterogeneous, and for any given citizen the likelihood of experiencing food insecurity will also be directly influenced by transport mobility, the local availability of affordable food, the prevalence of neighbours experiencing economic deprivation, and the wider embeddedness of social support networks. The consequence of these factors is that any theoretical measure of food insecurity is likely to be incomplete unless it is sensitive to a range of covariates. The results demonstrate that food surplus sharing arrangements such as OLIO must be considered if any measure of food insecurity is to be comprehensive. Existing deprivation measures are static whilst the data suggests that acute food needs are temporally dynamic (often over quite short time frames and presumably relating to real income variations). A basic point is justified in the overall approach adopted here: a ‘transactional’ data-set is better placed to provide insight into and to predict something temporally and intrinsically dynamic. Static deprivation measures cannot compete.

5.3 Implications for practice and policy

Previous work (e.g. Loopstra and Tarasuk, 2015) has highlighted the limitations of partial or incomplete data collection when measuring food insecurity and the only solution to this issue would be to compile public and proprietary data simultaneously from behavioural records and self-reported survey data. We agree with the United Nations Special Rapporteur who suggested that “The UK should introduce a single measure of insecurity and measure food security” (2018, p.23). However, unless the organisations chiefly responsible for emergency food assistance (e.g. foodbanks, community caf´es, food sharing applications) compile shared aggregate figures the prevalence of food insecurity (and its relation to food surplus) will remain obscure.

The findings also reveal the managerial need for food surplus sharing initiatives such as OLIO to monitor the prevalence of food insecurity longitudinally. The causes of food insecurity are systemic to the economy and food sharing organisations are limited in the level of assistance or relief that can be provided for those in acute food insecurity. But there is nonetheless a moral requirement on behalf of sharing economy organisations to document this previously hidden population.

6. Limitations and future research

The ‘ground truth’ used in the analysis relies upon self-declared food insecurity. The consequence is that the prevalence estimate of users in food insecurity - 12.08% - may be lower than the actual prevalence, as some users experiencing food insecurity may be unwilling to broadcast their hardship to the network. However, given that the predictive model is based on multiple behavioural features, those users who might feel uncomfortable requesting food may nonetheless exhibit similar behavioural repertoires (e.g. an emphasis on requesting food), and would thus be recognised by the model. These figures, too, are in line with D. Smith, Thompson et al.’s (2018) 17.40% estimate of population-level risk of food insecurity in English neighbourhoods using public data. Future research could help to scrutinise the declarative behaviour of users experiencing food insecurity to improve the quality of the ground truth used in predictive approaches.

One further practical limitation for implementing the predictive model, is that the temporal analysis involved is not dynamic i.e. the model is not updated systematically. We suggest that a practical implementation of the model should incorporate dynamic events (e.g. users deleting their accounts, users being banned by OLIO, or users experiencing a change in socioeconomic status), to ensure and improve the ongoing efficacy of prediction. If a dynamic model is implemented, a further study could examine how users move in and out of food insecurity, and thus shed further light on the factors which conspire to cause this pernicious social problem.

References

Al Hasan, Mohammad, Chaoji, Vineet, Salem, Saeed and Zaki, Mohammed (2006). ‘Link Prediction Using Supervised Learning’. In: SDM’06: Workshop on Link Analysis, Counterterrorism and Security. Maryland, USA.

Alaimo, Katherine, Olson, Marie and Frongillo, Edward A. (2001). ‘Food Insufficiency and American School-aged Children’s Cognitive, Academic, and Psychosocial Development’. In: Paediatrics 108.1, pp. 44–53.

Anderson, Annie S., Dewar, John, Marshall, David, Cummins, Steven, Taylor, Mathew, Dawson, John and Sparks, Leigh (2007). ‘The Development of a Healthy Eating Indicator Shopping Basket Tool (HEISB) for Use in Food Access Studies—Identification of Key Food Items’. In: Public Health Nutrition 10.12, pp. 1440–1447.

Ashton, John R., Middleton, John and Lang, Tim (2014). ‘Open Letter to Prime Minister David Cameron on Food Poverty in the UK’. In: The Lancet 383.9929, p. 1631.

Baron, Steve, Patterson, Anthony, Maull, Roger and Warnaby, Gary (2018). ‘Feed People First: A Service Ecosystem Perspective on Innovative Food Waste Reduction’. In: Journal of Service Research 21.1, pp. 135–150.

Black, Christina, Ntani, Georgia, Inskip, Hazel, Cooper, Cyrus, Cummins, Steven, Moon, Graham and Baird, Janis (2014). ‘Measuring the Healthfulness of Food Retail Stores: Variations by Store Type and Neighbourhood Deprivation’. In: International Journal of Behavioral Nutrition and Physical Activity 11.69, pp. 1–11.

Boguñá, Marián, Pastor-Satorras, Romualdo, Díaz-Guilera, Albert and Arenas, Alex (2004). ‘Models of Social Networks Based on Social Distance Attachment’. In: Physical Review E 70.5, pp. 70–78.

Bonacich, Phillip (1972). ‘Factoring and Weighting Approaches to Status Scores and Clique Identification’. In: The Journal of Mathematical Society 2.1, pp. 113–120.

Boone, Derrick S. and Roehm, Michelle (2002). ‘Retail Segmentation Using Artificial Neural Networks’. In: International Journal of Research in Marketing 19.3, pp. 287–301. Borgatti, Stephen P. (2006). ‘Identifying Sets of Key players in a Social Network’. In: Computational & Mathematical Organization Theory 12.1, pp. 21–34.

Borgatti, Stephen P., Jones, Candace and Everett, Martin G. (1998). ‘Network Measures of Social Capital’. In: Connections 21.2, pp. 27–36.

Botsman, Rachel and Rogers, Roo (2010). What’s Mine Is Yours: How Collaborative Consumption is Changing the Way We Live. Harper Collins Business.

Breitsohl, Jan, Kunz, Werner H. and Dowell, David (2015). ‘‘Does the Host Match the Content?’ A Taxonomical Update on Online Consumption Communities’. In: Journal of Marketing Management 31.9-10, pp. 1040–1064.

Brynjolfsson, Erik, Hitt, Lorin M. and Kim, Heekyung Hellen (2011). ‘Strength in Numbers: How Does Data-driven Decision Making Affect Firm Performance?’ In: SSRN. URL: https://ssrn.com/abstract=1819486 (visited on 03/02/2019).

Bucher, Eliane, Fieseler, Christian and Lutz, Christoph (2016). ‘‘What’s Mine Is Yours (for a Nominal Fee)’ — Exploring the Spectrum of Utilitarian to Altruistic Motives for Internet- Mediated Sharing’. In: Computers in Human Behavior 62, pp. 316–326.

Caplan, Pat (2017). ‘Win-Win? Food Poverty, Food Aid and Food Surplus in the UK Today’. In: Anthropology Today 33.3, pp. 17–22.

Caraher, Martin and Furey, Sinéad (2017). ‘Is it Appropriate to Use Surplus Food to Feed People in Hunger? Short term Band-Aid to More Deep-Rooted Problems of Poverty’. In: Food Research Collaboration Briefing Paper.

Carrigan, Marylyn (2017). ‘Revisiting ‘The Myth of the Ethical Consumer’: Why Are We Still Not Ethical Shoppers?’ In: Journal of Consumer Ethics 1.1, pp. 11–21.

Chaki, Jyotismita and Dey, Nilanjan (2019). ‘Pattern Analysis of Genetics and Genomics: A Survey of the State-of-art’. In: Multimedia Tools and Applications, pp. 1–32.

Cialdini, Robert and Schroeder, David A. (1976). ‘Increasing Compliance by Legitimizing Paltry Contributions: When even a Penny Helps’. In: Journal of Personality and Social Psychology 34, pp. 599–604.

Clary, Christelle, Lewis, Daniel J., Flint, Ellen, Smith, Neil R., Kestens, Yan and Cummins, Steven (2010). ‘Practice of Epidemiology the Local Food Environment and Fruit and Vegetable Intake: A Geographically Weighted Regression Approach in the ORiEL Study’. In: American Journal of Epidemiology 184.11, pp. 837–846.

Credit Suisse (2019). Global Wealth Report 2019. URL: https://www.credit-suisse.com/articles/media-releases/2019/10/en/global-wealth-report-2019--global-wealth-rises-by-2-6--driven-by.html (visited on 05/01/2020).

Cummins, Steven, Smith, Dianna, Aitken, Zoe, Dawson, John, Marshall, David, Sparks, Leigh and Anderson, Annie S. (2010). ‘Neighbourhood Deprivation and the Price and Availability of Fruit and Vegetables in Scotland’. In: Journal of Human Nutrition and Dietetics 23.5, pp. 494–501.

Cummins, Steven, Smith, Dianna, Taylor, Mathew, Dawson, John, Marshall, David, Sparks, Leigh and Anderson, Annie S. (2009). ‘Variations in Fresh Fruit and Vegetable Quality by Store Type, Urban – Rural Setting and Neighbourhood Deprivation in Scotland’. In: Public Health Nutrition 12.11, pp. 2044–2050.

Department for Communities and Local Government (2015). English Indices of Deprivation 2015. URL: https://www.gov.uk/government/statistics/englishindices-of-deprivation-2015 (visited on 01/12/2015).

Department for Work and Pensions (2017). DWP Benefit Statistics Tools. URL: https://www.gov.uk/guidance/dwp-benefit-statistics-disseminationtools (visited on 15/03/2019).

DiMaggio, Paul, Nag, Manish and Blei, David (2013). 'Exploiting Affinities between Topic Modeling and the Sociological Perspective on Culture: Application to Newspaper Coverage of US Government Arts Funding'. In: Poetics 41.6, pp. 570–606.

Dowler, Elizabeth and Lambie-Mumford, Hannah (2015a). 'Food for Thought: An Ethnographic Study of Negotiating Ill Health and Food Insecurity in a UK Foodbanks'. In: Social Policy and Medicine 132, pp. 38–44.

— (2015b). 'How can Households Eat in Austerity? Challenges for social policy in the UK'. In: Social Policy and Society 14.3, pp. 417–428.

Ducruet, César and Beauguitte, Laurent (2014). 'Spatial Science and Network Science: Review and Outcomes of a Complex Relationship'. In: Networks and Spatial Economics 14.3-4, pp. 297–316.

Eagle, Nathan and Pentland, Alex (2009). 'Eigenbehaviors: Identifying Structure in Routine'. In: Behavioral Ecology and Sociobiology 63.7, pp. 1057–1066.

Ehrenberg, Andrew S. C. (1988). Repeat-Buying: Facts, Theory, and Applications.

Engelmann, Gregor, Goulding, James and Smith, Gavin (2018). 'The Unbanked and Poverty: Predicting Area-level Socio-Economic Status from M-money Transactions'. In: IEEE Big Data. IEEE. Seattle, USA.

Enloe, James (2003). 'Food Sharing Past and Present: Archaeological Evidence for Economic and Social Interactions'. In: Before Farming 1, pp. 1–23.

FAO, IFAD, UNICEF, WFP and WHO (2017). The State of Food Security and Nutrition in the World 2017. Building Resilience for Peace and Food Security. Rome: FAO.

Foedermayr, Eva K. and Diamantopoulos, Adamantios (2008). 'Market Segmentation in Practice: Review of Empirical Studies, Methodological Assessment, and Agenda for Future Research'. In: Journal of Strategic Marketing 16.3, pp. 223–265.

Food Standards Agency (2019). UK Food Hygiene Rating Data API. URL: http://ratings.food.gov.uk/open-data/en-GB (visited on 16/01/2019).

Freeman, Linton C. (1979). 'Centrality in Social Networks: Conceptual Clarification'. In: Social Networks 1.3, pp. 215–239.

Galli, Francesca, Cavicchi, Alessio and Brunori, Gianluca (2019). 'Food Waste Reduction and Food Poverty Alleviation: A System Dynamics Conceptual Model'. In: Agriculture and Human Values, pp. 1–12.

Gautam, Bishnu Prasad and Shrestha, Dipesh (2010). ‘Document Clustering Through Nonnegative Matrix Factorization: A Case Study of Hadoop for Computational Time Reduction of Large-Scale Documents’. In: International Multi-Conference of Engineers and Computer Scientists IMECS 2010. Vol. I. Hong Kong, China.

George, Gerard, Haas, Martine R. and Pentland, Alex (2014). ‘From the Editors: Big Data and Management’. In: Academy of Management Journal 57.2, pp. 321–326.

GOV.UK (2018). National Public Transport Access Nodes. URL: https://www.gov.uk/government/publications/national-public-transport-accessnode-schema (visited on 16/11/2018).

Gustavsson, Jenny, Cederberg, Cristel and Sonesson, Ulf (2011). Global Food Losses and Food Waste. Rome, Italy.

Hackett, Michelle, Melgar-Quinonez, Hugo and Alvarez, Martha C. (2009). ‘Household Food Insecurity Associated with Stunting and Underweight among Preeschool Children in Antioquia, Colombia’. In: Revista Panamericana de Salud Publica 25, pp. 506–510.

Hagberg, Aric, Swart, Pieter and S. Chult, Daniel (2008). ‘Exploring Network Structure, Dynamics, and Function Using Networkx’. In: OSTI.GOV. URL: https://permalink.lanl.gov/object/tr?what=info:lanl-repo/lareport/LA-UR-08-05495 (visited on 20/12/2018).

Han, Jiawei, Pei, Jian and Kamber, Micheline (2011). Data Mining: Concepts and Techniques. Massachusetts, USA: Elsevier.

Harvey, John, Smith, Andrew, Goulding, James and Branco-Illodo, Ines (2019). ‘Food Sharing, Redistribution, and Waste Reduction via Mobile Applications: A Social Network Analysis’. In: Industrial Marketing Management Forthcoming.

Hellwig, Katharina F., Morhart, Felicitas, Girardin, Florent and Hauser, Mirjam (2015). ‘Exploring Different Type of Sharing: A Proposed Segmentation of the Market for ‘Sharing’ Businesses’. In: Psychology and Marketing 32.9, pp. 891–906.

Holme, Petter and Saramäki, Jari (2013). Temporal Networks. Springer.

Hopcroft, John, Lou, Tiancheng and Tang, Jie (2011). ‘Who Will Follow you Back?: Reciprocal Relationship Prediction’. In: Proceedings of the 20th ACM International Conference on Information and Knowledge Management. Glasgow, Scotland, pp. 1137–1146.

House of Commons Environment, Food and Rural Affairs Committee (2017). Food Waste in England, Eighth Report of Session 2016-17. URL: https://publications.parliament.uk/pa/cm201617/cmselect/cmenvfru/429/429.pdf (visited on 01/12/2018).

Hruschka, Harald (2014). ‘Linking Multi-category Purchases to Latent Activities of Shoppers: Analysing Market Baskets by Topic Models’. In: Marketing: ZFP–Journal of Research and Management 36.4, pp. 267–273.

Iijima, Ryota and Kamada, Yuichiro (2017). ‘Social Distance and Network Structures’. In: Theoretical Economics 12.2, pp. 655–689.

Independent Food Aid Network (2018). Mapping the UK’s Independent Food Banks. URL: http://www.foodaidnetwork.org.uk/mapping (visited on 16/11/2018).

Jaeggi, Adrian V. and Gurven, Michale (2013). ‘Natural Co-operators: Food Sharing in Humans and Other Primates’. In: Evolutionary Anthropology: Issues, News, and Reviews 22.4, pp. 186–195.

Kubina, Milan, Varmus, Michal and Kubinova, Irena (2015). ‘Use of Big Data for Competitive Advantage of Company’. In: Procedia Economics and Finance 26, pp. 561–565.

Lamberton, Cait Poynor and Rose, Randall L. (2012). ‘When is Yours Better than Mine? A Framework for Understanding and Altering Participation in Commercial Sharing Systems’. In: Journal of Marketing 76.4, pp. 109–125.

Lazell, Jordon, Magrizos, Solon and Carrigan, Marylyn (2018). ‘Over-claiming the Circular Economy: The Missing Dimensions’. In: Social Business 8.1, pp. 103–114.

Lazer, David, Pentland, Alex, Adamic, Lada, Aral, Sinan, Barabási, Albert-László, Brewer, Devon, Christakis, Nicholas, Contractor, Noshir, Fowler, James, Gutmann, Myron et al. (2009). ‘Computational Social Science’. In: Science 323.5915, pp. 721–723.

Lee, Daniel D. and Seung, H. Sebastian (1999). ‘Learning the Parts of Objects by Non-negative Matrix Factorization’. In: Nature 401.6755, pp. 788–791.

Lilien, Gary L. and Rangaswamy, Arvind (2004). Marketing Engineering: Computer-assisted Marketing Analysis and Planning. DecisionPro.

London Datastore (2019). The First Food Security Estimates for London. URL: https://data.london.gov.uk/blog/the-first-food-security-estimatesfor-london/ (visited on 05/01/2020).

Loopstra, Rachel (2020). ‘Food Insecurity Measurement on the Family Resources Survey’. In: ENUF: Evidence and Network on UK Household Food Insecurity. URL: https://enuf.org.uk/research- blogs/food- insecurity- measurement-familyresources-survey (visited on 05/01/2020).

Loopstra, Rachel, Reeves, Aaron and Tarasuk, Valerie (2019). ‘The Rise of Hunger among Low-income Households: An Analysis of the Risks of Food Insecurity between 2004 and 2016 in a Population-based Study of UK Adults’. In: Journal of Epidemiology and Community Health 73, pp. 668–673.

Loopstra, Rachel, Reeves, Aaron, Taylor-Robinson, David, Barr, Ben, McKee, Martin and Stuckler, David (2015). 'Austerity, Sanctions, and the Rise of Food Banks in the UK'. In: British Medical Journal 350, p.h1775.

Loopstra, Rachel and Tarasuk, Valerie (2015). 'Food Bank Usage is a Poor Indicator of Food Insecurity: Insights from Canada'. In: Social Policy and Society 14.3, pp. 443–455.

McEntee, Jesse (2009). 'Highlighting Food Inadequacies: Does the Food Desert Metaphor Help This Cause?' In: British Food Journal 111.4, pp. 349–363.

McPherson, Miller, Smith-Lovin, Lynn and Cook, James M. (2001). 'Birds of a Feather: Homophily in Social Networks'. In: Annual Review of Sociology 27.1, pp. 415–444.

Michelini, Laura, Principato, Ludovica and Iasevoli, Gennaro (2018). 'Understanding Food Sharing Models to Tackle Sustainability Challenges'. In: Ecological Economics 145, pp. 205–217.

Moore, David J. and Harris, William D. (1996). 'Affect Intensity and the Consumer's Attitude toward High Impact Emotional Advertising Appeals'. In: Journal of Advertising 25.2, pp. 37–50.

NatCen Social Research (2017). Food and You Survey Wave 4 (2016). London, UK.
Nelson, Michael, Erens, Bob, Bates, Beverly, Church, Susan and Boshier, Tracy (2007). Low Income Diet and Nutrition Survey. Vol. 3. London, UK: TSO London.

Ng, Irene C. L., Maull, Roger S., Parry, Glenn C., Crowcroft, Jon, Scharf, Kimberley, Rodden, Tom and Speed, Chris (2013). 'Making Value Creation Context Visible for New Economic and Business Models: Home Hub-of-all-Things (H.A.T.) as Platform for Multi-sided Market Powered by Internet-of-Things'. In: Hawaii International Conference on Systems Science (HICSS) 2013. Hawaii, USA, pp. 7–10.

Northern Ireland Statistics and Research Agency (2017). NIMDM17-SOA levell results. URL: https://www.nisra.gov.uk/publications/nimdm17-soa-levelresults (visited on 01/12/2018).

OLIO (2019). OLIO - The Food Sharing Revolution. URL: https://olioex.com/ (visited on 06/03/2019).

Raschka, Sebastian (2018). 'Model Evaluation, Model Selection, and Algorithm Selection in Machine Learning'. In: arXiv preprint arXiv:1811.12808.

Ravasz, Erzsébet and Barabási, Albert-László (2003). 'Hierarchical Organization in Complex Networks'. In: Physical Review E 67.2, pp. 1–7.

Saito, Takaya and Rehmsmeier, Marc (2015). 'The Precision-Recall Plot Is More Informative than the ROC Plot when Evaluating Binary Classifiers on Imbalanced Datasets'. In: PloS one 10.3, e0118432.

Savage, Mike and Burrows, Roger (2007). ‘The Coming Crisis of Empirical Sociology’. In: Sociology 41.5, pp. 885–899.

Scaraboto, Daiane (2015). ‘Selling, Sharing, and Everything in Between: The Hybrid Economies of Collaborative Networks’. In: Journal of Consumer Research 42.1, pp. 152–176.

Schanes, Karin, Dobernig, Karin and Gozet, Burcu (2018). ‘Food Waste Matters- A Systematic Review of Household Food Waste Practices and their Policy Implications’. In: Journal of Cleaner Production 182, pp. 978–991.

Schanes, Karin and Stagl, Sigrid (2019). ‘Food Waste Fighters: What Motivates People to Engage in Food Sharing?’ In: Journal of Cleaner Production 211, pp. 1491–1501.

Schor, J. and Fitzmaurice, C. (2015). ‘Collaborating and Connecting: The Emergence of the Sharing Economy’. In: Handbook of Research on Sustainable Consumption. Ed. by Lucia Reisch and John Thogersen. Cheltenham, UK: Edward Elgar, pp. 410–425.

Scottish Government (2016). The Scottish Index of Multiple Deprivation. URL: https://www2.gov.scot/Topics/Statistics/SIMD (visited on 01/12/2018).

Shah, Denish, Kumar, V. and Kim, Kihyun Hannah (2014). ‘Managing Customer Profits: The Power of Habits’. In: Journal of Marketing Research 51.6, pp. 726–741.

Shearman, Sachiyo M. and Yoo, Jina H. (2007). ‘“Even a Penny Will Help!”: Legitimization of Paltry Donation and Social Proof in Soliciting Donation to a Charitable Organization’. In: Physical Review E - Statistical, Nonlinear, and Soft Matter Physics 24.4, pp. 271–282.

Sivarajah, Uthayasankar, Kamal, M. M., Irani, Z. and Weerakkody, V. (2017). ‘Critical Analysis of Big Data Challenges and Analytical Methods’. In: Journal of Business Research 70, pp. 263–286.

Smith, Dianna, Cummins, Steven, Taylor, Mathew, Dawson, John, Marshall, David, Sparks, Leigh and Anderson, Annie S. (2010). ‘Neighbourhood Food Environment and Area Deprivation: Spatial Accessibility to Grocery Stores Selling Fresh Fruit and Vegetables in Urban and Rural Settings’. In: International Journal of Epidemiology 39.1, pp. 277–284.

Smith, Dianna, Thompson, Claire, Harland, Kirk, Parker, Storm and Shelton, Nicola (2018). ‘Identifying Populations and Areas at Greatest Risk of Household Food Insecurity in England’. In: Applied Geography 91, pp. 21–31.

StatsWales (2014). Indicator Data by Lower Layer Super Output Areas - All Domains. URL: https://gov.wales/statistics-and-research (visited on 01/12/2018).

Strong, Colin (2015). Humanizing Big Data: Marketing at the Meeting of Data, Social Science and Consumer Insight. London, UK: Kogan Page Publishers.

Sun, Yanmin, Wong, Andrew K. C. and Kamel, Mohamed S. (2009). ‘Classification of Imbalanced Data: A Review’. In: International Journal of Pattern Recognition and Artificial Intelligence 23.04, pp. 687–719.

Tarasuk, Valerie, St-Germain, Andre/’e-Anne and Loopstra, Rachel (2019). ‘The Relationship between Food Banks and Food Insecurity: Insights from Canada’. In: VOLUNTAS: International Society for Third-Sector Research.

Taylor-Robinson, David, Rougeaux, Emeline, Harrison, Dominic, Whitehead, Margaret, Barr, Ben and Pearce, Anna (2013). ‘The Rise of Food Poverty in the UK’. In: British Medical Journal 347, f7157.

The Social Market Foundation (2018). Can Everyone Access Affordable, Nutritious Food? A Picture of Britain’s Deprived Food Deserts, pp. 1–10. URL: https://www.kelloggs.co.uk/content/dam/europe/kelloggs_gb/pdf/Kelloggs_Food_Desert_Brochure.pdf (visited on 19/02/2019).

The Trussell Trust (2018). Find a Foodbank. URL: https://www.trusselltrust.org/get-help/find-a-foodbank/ (visited on 16/11/2018).

Thompson, Claire, Smith, Dianna and Cummins, Steven (2018). ‘Understanding the Health and Wellbeing Challenges of the Food Banking System: A Qualitative Study of Food Bank Users, Providers and Referrers in London’. In: Social Science & Medicine 211, pp. 95–101.

— (2019). ‘Food Banking and Emergency Food Aid: Expanding the Definition of Local Food Environments and Systems’. In: International Journal of Behavioral Nutrition and Physical Activity 16.1, p. 2.

United Nations Development Program (2016). Sustainable Development Goals. URL: https://www.undp.org/content/undp/en/home/sustainable-developmentgoals.html (visited on 01/12/2019).

United Nations Special Rapporteur (2019). Statement on Visit to the United Kingdom, by Professor Philip Alston, United Nations Rapporteur on Extreme Poverty and Human Rights. URL: https://www.ohchr.org/en/NewsEvents/Pages/DisplayNews.aspx?NewsID=23881 (visited on 06/03/2019).

United States Department of Agriculture Economic Research Service (2017). Survey Tools. URL: https://www.ers.usda.gov/topics/food-nutrition-assistance/food-security-in-the-us/survey-tools.aspx (visited on 05/01/2020).

Vabalas, Andrius, Gowen, Emma, Poliakoff, Ellen and Casson, Alexander J. (2019). ‘Machine Learning Algorithm Validation with a Limited Sample Size’. In: PloS one 14.11.

Walker, Peter (2017). ‘Dominic Raab Accused of ‘Stupid and Offensive’ Food Bank Comments’. In: The Guardian. URL:

https://www.theguardian.com/politics/2017/may/29/tory-mp-dominic-raab-jeered-over-food-bankcomments (visited on 06/03/2019).

Whitaker, Robert C., Philips, Shannon M. and Orzol, Sean M. (2006). ‘Food Insecurity and the Risks of Depression and Anxiety in Mothers and Behavior Problems in their Preschool aged Children’. In: Paediatrics 113.3, e859–e868.

White, Katherine and Peloza, John (2009). ‘Self-benefit versus Other-benefit Marketing Appeals: Their Effectiveness in Generating Charitable Support’. In: Journal of Marketing 73.4, pp. 109–124.

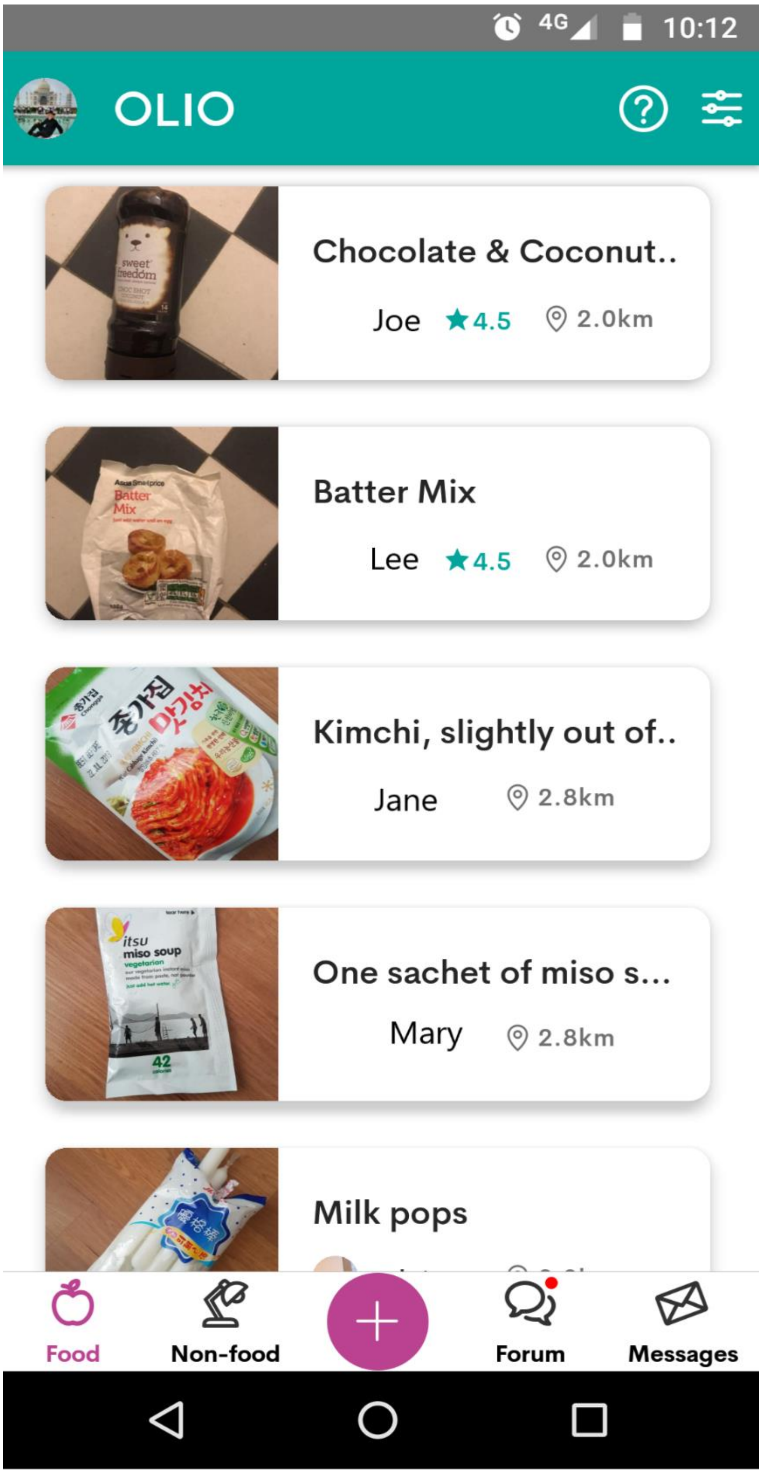

Figure 1: Redacted screenshot of OLIO application showing nearby food available for request.

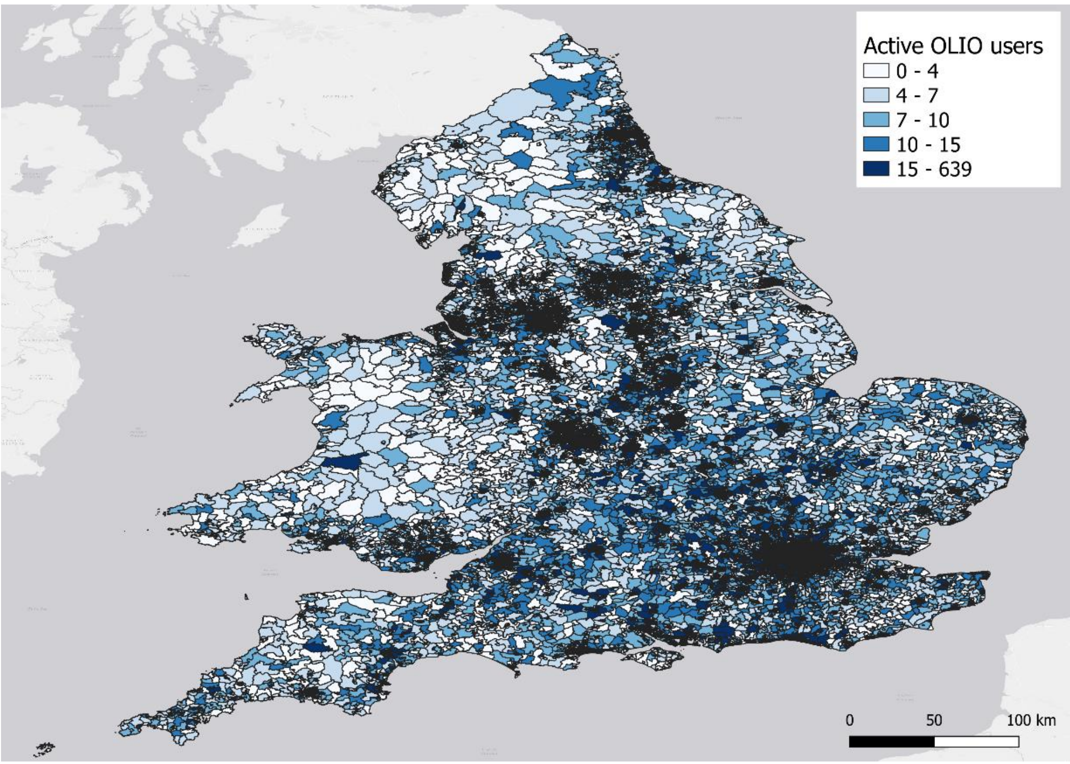

Figure 2: OLIO user densities in UK neighbourhoods (Wales and England visualised here). Olio users are present in one third of Britain's 'neighbourhoods', with most having between 1 and 15 active users.

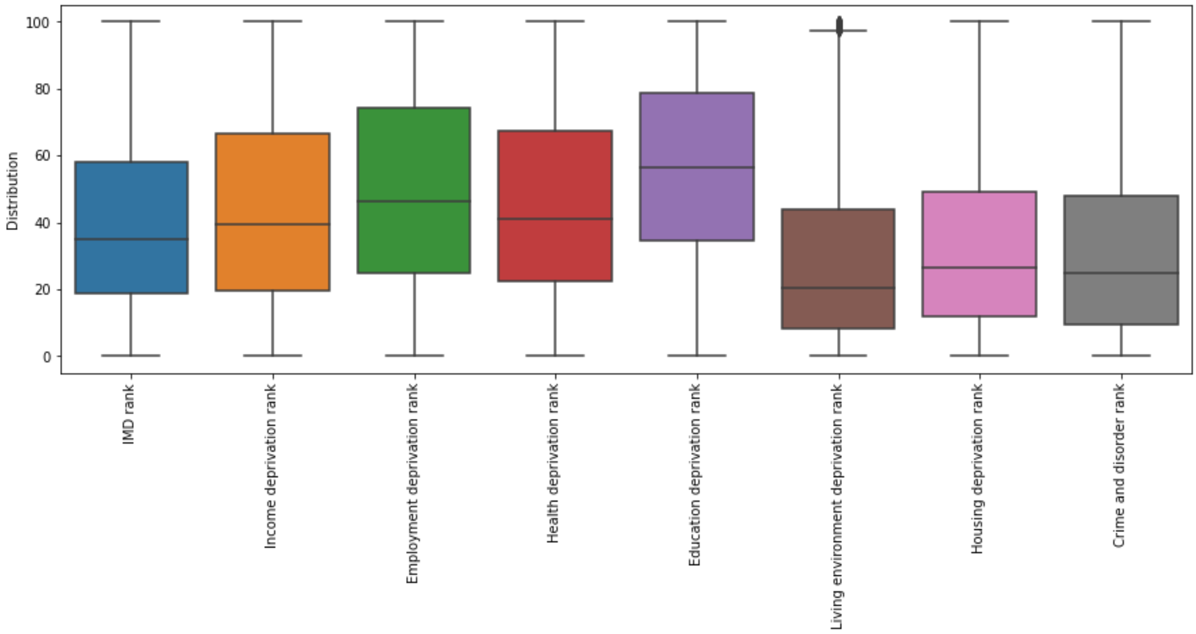

Figure 3: Composite measure and dimensions of deprivation characteristic of users' neighbourhood of registration. The average OLIO user lives in moderately deprived neighbourhoods - M=41.09, SD=28.48, with above average levels of education - M=55.57, SD=27.32, but poor living environments and housing - M=29.54, SD=26.84.

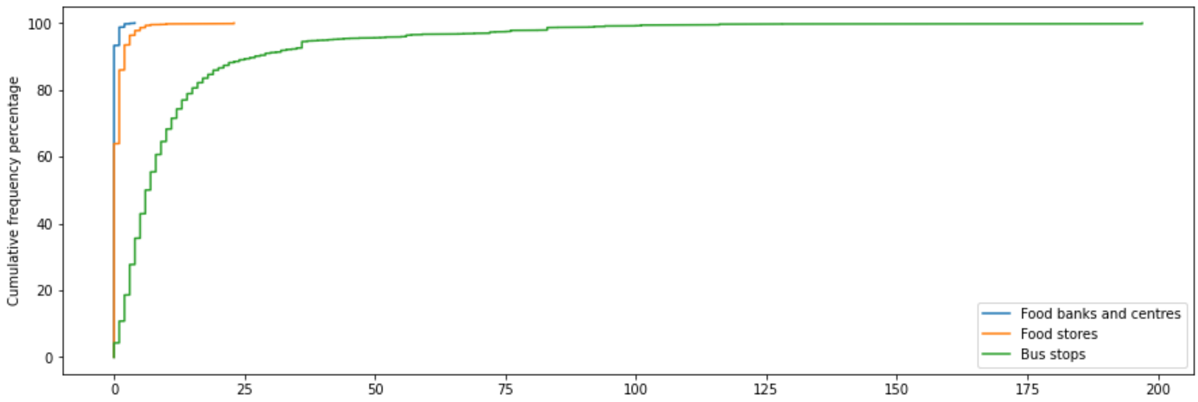

Figure 4: Number of locales in users' neighbourhoods. 93.31% of OLIO users in the UK have no food bank or centre in their vicinity, with 6% having between one and four. In terms of food stores, 63.90% of users have none in their neighbourhood, with a further 29.92% having between one and two. As far as public transport accessibility is concerned, 10% of OLIO users have at most one bus stop in their vicinity, with 68.27% having as many as 10, with the highest public transport density being recorded in central London.

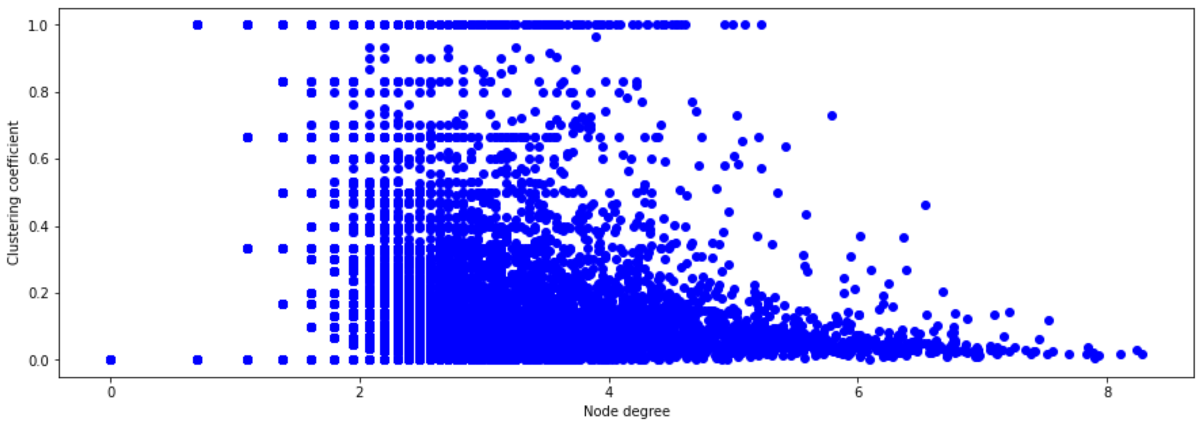

Figure 5: OLIO users' average clustering coefficient as a function of degree, on the natural logarithmic scale.

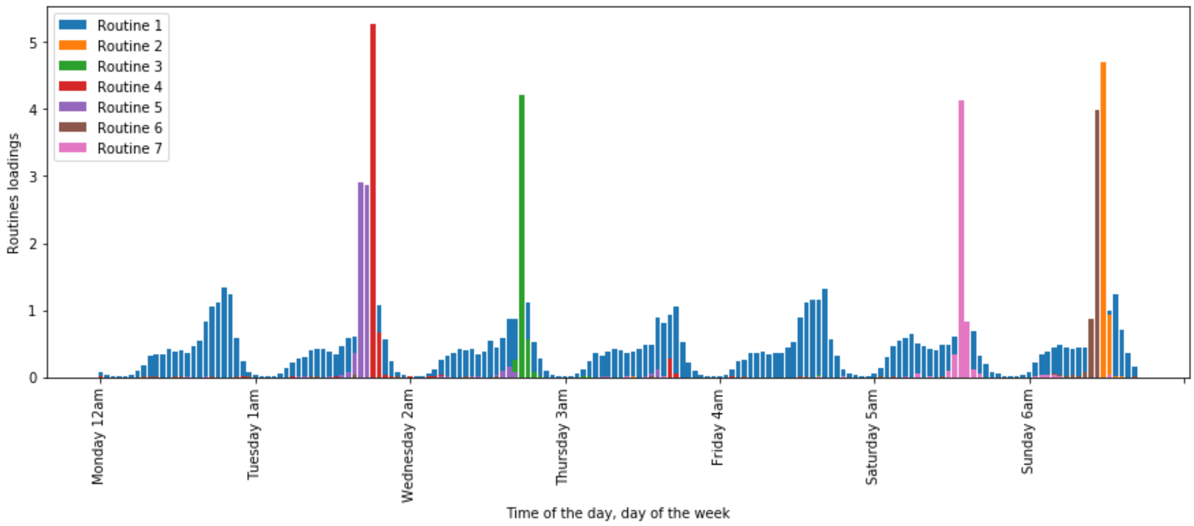

Figure 6: OLIO users' request routines as evidenced from transactional records. NMF was applied to time-stamped solicitation messages to identify latent routines across the sample. Seven weekly routines emerged: Routine 1, for example, reveals soliciting behaviour throughout the week, while Routines 2 and 6 emphasise weekend afternoons.

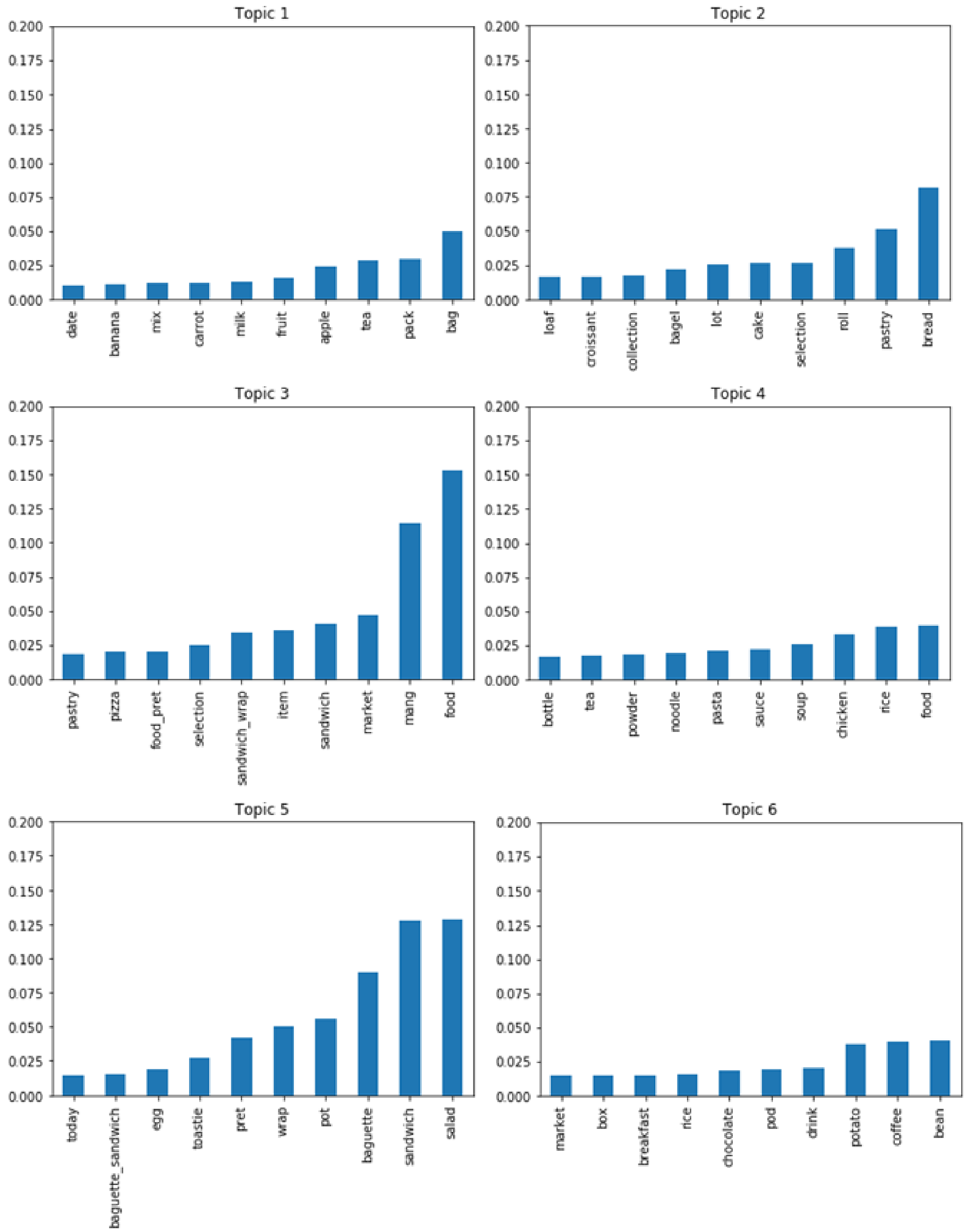

Figure 7: OLIO users' diet preferences as evidenced from transactional records. LDA was applied to corpus of soli-citation messages to identify latent diet preferences (topics) across the sample. Six such topics emerged: Topic 1, for example, reveals preference for fresh fruit and produce, Topic 2 - bakery products, Topics 3 and 5 emphasise sandwich and deli products sourced from associated stores, while Topic 4 - long life products. Any one OLIO user may be described as a combination of these diet preferences, and so may express each of these topics to a different amount.

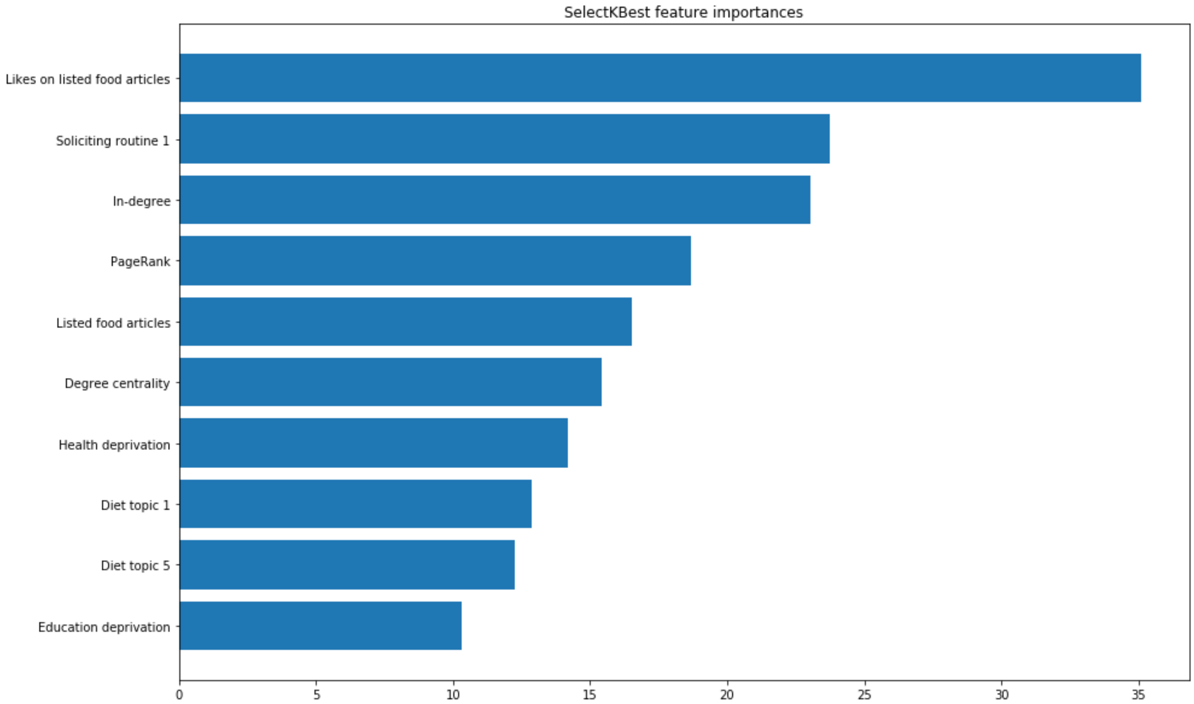

Figure 8: The ranking of features contributing to the prediction of one's self-declared food insecurity status, chosen in conjunction with Random Forests classifier.

Table 1: Dimensions of individual food insecurity

Dimensions of individual food insecurity				
Dimension	**Data**	**Source**	**Time frame**	**Description**
Neighbourhood characteristics	Normalised composite measure and domains of deprivation, including employment, income, health, education, living environment, crime and housing	English Indices of Deprivation (IMD 2015, Department for Communities and Local Government, 2015)	2015 measure based on 2012-2013 data	Ranking of 32,844 Lower Super Output Areas (LSOAs) with an average of 1,500 inhabitants based on composite measure and domains of deprivation.
		Northern Ireland Multiple Deprivation Measure (NIMDM 2017, Northern Ireland Statistics and Research Agency, 2017)	2017 measure based on 2016 data	The NIMDM measures deprivation across seven domains that produce an overall rank of 890 Super Output Areas (SOAs) averaging 2,000 inhabitants.
		Welsh Multiple Deprivation Measure (WIMD 2014, StatsWales, 2014)	2014 measure updated in 2015	The WIMD is used to rank each of the 1,909 LSOAs in Wales with an average population of 1,600 from 1 (most deprived) to 1,909 (least deprived).
		Scottish Indices of Deprivation (SIMD 2016, Scottish Government, 2016)	2016 measure	The SIMD ranks each of Scotland's 6,976 Data Zones (DZ) of between 500 and 1,000 households.
	Proportion of benefit claimants in UK neighbourhoods	Employment & Support Allowance (ESA) and Pension Credit (PC) (Department for Work and Pensions, 2017)	May- August 2018	Number of people claiming benefits by LSOA, normalised by 2011 population census.
	Bus stop locations	National Public Transport Access Node Schema (NaPTAN GOV.UK, 2018)	Published in 2014, updated in 2017	Location of approximately 400,000 bus stops in the UK, with locality and latitude/ longitude details
	Foodbank locations	Trussell Trust Network (The Trussell Trust, 2018)	1,707 locations scraped in November 2018 using Python APIs	The Trussell Trust runs a network of foodbanks and centres across the UK and provides emergency food to people referred for support.

		Independent Food Aid Network (IFAN, Independent Food Aid Network, 2018)	653 locations scraped in November 2018 using Python APIs	Since 2017 the IFAN has been documenting the location of independent food banks in the UK.
	Food store locations	Food Standards Agency (FSA, Food Standards Agency, 2019)	12,009 food stores locations across the UK were scraped in March 2019	This platform provides data on food hygiene ratings or inspection results for businesses including restaurants, pubs, cafés, takeaways, hotels and other places where consumers eat, as well as supermarkets and other food stores.
Network topology	OLIO platform data	OLIO proprietary data	The transactional data contains a three-year period from 9th July 2015 until 30th October 2018	Dataset can be best summarised as conversations among users, whereby some offered items, others requested them, with food exchanged offline being marked as successfully ‘picked up’. These conversations also include a category of ‘wanted items’ - immediate and specific requests.
Behavioural repertoires	OLIO platform data	OLIO proprietary data	The transactional data contains a three-year period from 9th July 2015 until 30th October 2018	There are 141,129 unique items listed in the dataset, of which 102,239 were requested 238,622 times and 99,604 were exchanged offline by 41,811 users of the network in the UK.

Table 2: Behavioural repertoires on the OLIO platform. Three main behaviours emerged, donating, requesting and collecting food from associated stores for redistribution into the wider community, with varying levels of reciprocation

Behavioural repertoires on the OLIO platform		
Network role	**Sample size**	**Percentage**
Active users on the OLIO platform	**41,811**	**100%**
Donors	**17,172**	**41.07%**
Donors who have not requested	11,685	27.95%
Donors who have requested and received food	2,579	6.17 %
Donors who have requested but not received food	2,311	5.53%
Users who have requested	**30,065**	**71.90%**
Users whose requests were met	11,093	26.53%
Users whose requests were not met	13,402	32.05%
Volunteers collecting food for redistribution	**741**	**1.78%**
Volunteers, donors, requesters and receivers of food	523	1.27%
Volunteers, donors, requesters	114	0.27%
Volunteers, requesters and receivers of food	104	0.70%

Table 3: Dimensions of food insecurity

Dimensions of food insecurity						
Dimension	**Neighbourhood characteristics**		**Network topology**		**Behavioural repertoires**	
		↓		↘		↙
	National statistics aggregated at neighbourhood level (LSOA)				Observed food sharing behaviours on OLIO	
		↓		↓		↓
Target data	IMD rank	Aggregate statistics	In-degree	Network analysis	Food soliciting routines	NMF
	Income deprivation rank		Out-degree		Diet preferences topics	LDA
	Employment deprivation rank		Betweenness centrality		Recency, frequency and quantity of requests	Aggregate statistics
	Health deprivation rank		Closeness centrality		Liked articles	
	Education deprivation rank		Degree centrality		Likes received on listed articles	
	Access to services deprivation rank		PageRank		Collections	
	Living environment deprivation rank		Clustering coefficient		Listed food articles	
	Crime rank				Average exchange distance	
	Probability of soliciting benefits				Maximum number of requests per hour	
	Number of bus stops				Maximum number of requests per day	
	Distance to nearest bus stop				Maximum number of requests per month	
	Number of food stores				Daily soliciting burstiness	
	Distance to nearest food store					
	Number of foodbanks					
	Distance to nearest foodbank			↓		
Results (Q1)	Comparison of aggregate statistics: correlation analysis					
			↓			
Induction (RQ2)	Explanatory account of food insecurity: feature ranking and selection; classification					

Table 4: For each model, the data was split into a training (four-fifths) and test set (one-fifth) and the parameters for the four classifiers (e.g., number of trees for RF, nearest neighbour for kNN), as well as for each feature selection approach (number of components for PCA and number of best predictive features for feature ranking) selected via a grid search underpinned by five fold cross-validation.

Classifier descriptions and parameters		
Classifier	**Description**	**Parameters**
Random Forests (RF)	Random Forests is a collection of decision tree classifiers fitted on subsamples of the data set, whereby the class (food secure of insecure) is predicted by popular vote across the trees (Han, Pei and Kamber, 2011).	• Estimators (number of trees in the forest): range 5- 100 • Number of components/ best predictive features: range 5- 45
Adaptive Boosting (AdaBoost)	AdaBoost is another ensemble algorithm, whose accuracy of classification is adjusted by adapting the weights of incorrectly classified instances (Han, Pei and Kamber, 2011).	• Estimators (number of estimators at which weight boosting is terminated): range 5- 100 • Number of components/ best predictive features: range 5- 45
Support Vector Machines	A support vector machine transforms the feature space into a higher dimension, 'where it finds a hyperplane that separates the data by class' (Han, Pei and Kamber, 2011, p.393).	• Kernel types to be used in the algorithm: linear and radial basis function (RBF) • Number of components/ best predictive features: range 5- 45
k-Nearest-Neighbors (kNN)	A user is classified as food secure or insecure based on the class of its k most similar users (or neighbours) (Han, Pei and Kamber, 2011).	• Number of neighbours to consider: 3, 5, 7 • Number of components/ best predictive features: range 5- 45

Table 5: Predicting instances of food insecurity. The performance of each model was measured via the average accuracy, precision and recall scores across five runs per model. The RF and AdaBoost models followed by ranking and selection of top features performed best. For the former, out of all the users classified as experiencing food insecurity, 76% were correct (precision), with 75% of all the instances of food insecurity in the sample being classified as such (recall).

Grid search of PCA and classifiers				
PCA *n* components	**Classifier**	**Accuracy**	**Precision**	**Recall**
40 estimators, 20 components	RF	58.82%	0.59	0.59
80 estimators, 30 components	AdaBoost	63.53%	0.64	0.64
RBF kernel, 15 components	SVM	63.53%	0.64	0.64
7 neighbours, 15 components	kNN	65.88%	0.66	0.66
Grid search of top- ranking features and classifiers				
75 estimators, 10 features	RF	75.29%	0.76	0.75
20 estimators, 45 features	AdaBoost	71.76%	0.72	0.72
RBF kernel, 30 features	SVM	63.53%	0.63	0.64
3 neighbours, 30 features	kNN	56.47%	0.56	0.56