Omnifood – Exploring the Possibilities of a Consumer System With Ubiquitous Access to Data About the Food We Eat

Björn Hedin
Department of learning
ITM/KTH Royal Institute of Technology
Stockholm, Sweden
bjornh@kth.se

ABSTRACT

Data about food, and data about individuals' purchases and consumption of food are becoming increasingly ubiquitous. Through bonus cards, supermarkets can track exactly which products we buy, through diet apps we can track what we eat, and through blockchains and other technologies producers can track the origin and history of individual products. From a technical point of view, we are not far away from a world where all this information could be combined to one omniscient system - OmniFood.

In this paper we explore current possibilities to collect data on what products we buy, how environmental and nutritional data can be mapped to these products and possibilities to track what we actually eat. Next, we present a number of prototype systems where the possibilities to use this data has been explored, and what limitations we have encountered with current implementations and available data. We end with a discussion of some services that could be possible if current technologies would be fully implemented and made available to consumers and system developers. What possibilities could such systems offer for consumers who want to eat both sustainable and healthy food? What limitations would still exist? What ethical aspects would need to be considered? The focus is on using such a system as a decision support system to support consumers in making food purchase choices that are sustainable from both environmental and health perspectives, thereby supporting the global food system to stay within sustainable limits.

KEYWORDS

Sustainable Food Systems, Sustainable HCI, Behavior Change, Digital Behavior Change, Consumption Data, Sustainability Data

ACM Reference format:

Björn Hedin. 2022. OmniFood – Exploring the possibilities of a consumer system with ubiquitous access to data about the food we eat. In *Proceedings Computing within limits, 2022, 9 pages.*

1 Introduction

The food system is one of the most essential components of today's society. Different choices in *how* food is grown and consumed, and *which* food is grown and consumed has high impact on several sustainability-related areas, including social, ecological and economical sustainability. Several hard and soft limits related to environmental sustainability exist such as land use, greenhouse gas emissions and use of fertilizers. Furthermore, the food system has several sustainability limits related to social and economical sustainability such as conditions for farmers and eating sufficiently nutritious food. Many of us want to maintain a healthy and affordable diet within social and ecological boundaries, but doing so is complex, not least since improvements in one area can adversely affect another area. For example, almonds are considered healthy and can even have negative emissions of greenhouse gases, but requires immense amounts of water to grow, ecological food is considered good for biodiversity but often require more arable land and so on. As a consumer it is difficult to know if a choice regarding what to eat is actually good or bad.

Until the digital era it has been very difficult as a consumer to make well-grounded decisions regarding healthy and sustainable food choices, where central problems include lack of data about production methods, transport, personal consumption, nutritional content, and personal nutritional needs. However, due to the digitalization of society, much of these data are now becoming available in small "data islands".

Nutritional values, much more detailed than what can be addressed on a label, are available in various databases on a product level, purchase data are collected by the food retailers and can in some cases be available for the consumer, and the origin and shipping history of products can be tracked through blockchains to give a few examples. These kinds of data however are not yet available in such a way that they can all be used together by an integrated system, and the full range of possible services are therefore not yet realized. However, there are also several good reasons for having a cautionary attitude towards integrating these data, including personal integrity, ethics, and business reasons.

In this paper, we start with exploring the current possibilities and limits for

- methods to collect purchase data
- methods to identify products
- environmental and nutritional data about products
- methods to identify what we eat and waste

Next, we present a number of prototype systems we have developed that use these methods and data, and the problems we have encountered. We end with a discussion of what kinds of services would be possible if all methods available were fully integrated into the food system, with a focus on decision support system to support consumers in making food purchase choices that are sustainable from both environmental and health perspectives, thereby supporting the global food system to stay within sustainable limits.

1.1 Problem Areas From a LIMITS Perspectiv

The global food system today is not sustainable and has several limits that need to be considered. These include for example CO_2 emissions, limited supply of fertilizers, limited water availability and limited availability av arable land. The consumers are key players in keeping the food system within these limits since they, by changing their food consumption behaviour, can reduce the stress on these limited resources.

1.1.1 Malnutrition and Overconsumption

Malnutrition comes in two forms. Undernutrition can be a problem in many countries. In the western world it is mainly a problem for the elderly, but for certain nutrients there can be problems for all groups. For example, in Sweden, 9 out of 10 young adults eat less fibers than the recommendation [8]. However, misconceptions are common, where for example many people worry about not getting enough protein, whereas that is very rarely a problem for most people [13]. Indeed, overnutrition and overconsumption is now a more common global health problem, where obesity and overweight caused by an increased consumption of energy-dense food leads to more deaths than undernutrition [14].

Both undernutrition and overnutrition are important from a health perspective, but from an individual's point of view it is not always easy to keep track on what you eat and if you get nutrients that are within the suggested limits. From an ecological sustainability point of view, overconsumption is the bigger problem, since eating more than what is healthy also means consuming more resources than necessary. Therefore, finding ways to support people to eat a healthy amount of healthy food is not just a solution to more a healthier population, but also a way to support a food system that stays within sustainable limits.

1.1.2 Food Waste

Food waste is also a problem related to overproduction of food. About one third of all the food produced for humans are wasted on a global level, where most of the food waste occurs in households. While there are several reasons for food waste in the household, one of the more important issues are a lack of food supply awareness, meaning that household members are unaware of what food they have already purchased, and therefor tend to buy more food than what is used. This results in unnecessary food waste when the food passes its expiry date. There have been several attempts to build personal food inventory systems, but they have so far always involved manual registration, which results in the users forgetting or lacking the energy to register and de-register food items, and the system quickly going out of use.

1.1.3 Resource-Inefficient Food

Another problem from a limits perspective is food products that require large amounts of resources to produce and distribute such as water, land, fuel and transportation. General awareness about the resources used for, and environmental impact by, food products are low, and choosing different kinds of food could substantially support a food system within sustainable limits.

1.1.4 Limits on Technical Systems

Finally, there are limits on what a decision support system for supporting sustainable food consumption can do and what it theoretically could do.

2 Current Enabling Technologies

In order to build system that support sustainable food using data, we propose that there are four main aspects that has to be considered; knowing what we actually buy, linking what we have bought to specific products, finding relevant data about these products and finally to know what we eat and what we waste.

2.1 Knowing What We Buy - Purchase Data

A central part of any system designed around analyzing the sustainability of food practices is to have access to grocery purchase data. There are several current solutions and possibilities to address this.

2.1.1 Extracting Data From Receipts

Some solutions have been designed around scanning and parsing paper receipts to extract data from these [12, 17, 19]. Such systems have two major problems. First, few customers would have the stamina to continue such a practice over an extended period of time, and second, to identify specific products and amounts bought can be difficult. Often the only identifier for a product is a short string that fits within a very limited number of characters, like "OLIVOLJA SPAN.EKO". A parser or a human might be able to guess that this is Spanish organic olive oil, but neither the specific brand nor how much oil it is can be deduced from that string. Even though this limited data can provide some valuable information, for example by general information about olive oil, it is of limited value. On the other hand, these receipts often contain weight on items bought in bulk such as fruit and vegetables, and for these it is relatively easy identify both the type of food bought ("apples", "chicken filet"), making it possible to identify environmental footprint and nutritional data based on model values for each type of food (see section 2.3.1). However, it is not possible to discern between for example avocado produced locally in Sweden from avocado produced in Peru.

Getting receipts in a digital format rather than a physical format provides several advantages. One argument often brought forward from the supermarkets chains is that digital receipts reduce paper use, but the main two benefits from the perspective of this paper are arguably that digital receipts can be automatically collected and processed, for example by companies handling expenses, and that they can contain much more information than can be found on a paper slip.

Some Swedish supermarket chains have the option to get the receipts in a digital format, but their methods vary. In section 3.1.1 we describe a prototype we have developed where the receipts are sent out by the supermarket chaihn by email immediately after payment. The system we designed can then automatically forward the receipt to a parser, which removes the need for manual processing/scanning of the receipts. However, these receipts are just digital versions of the physical receipt with exactly the same layout and, and thus provides no help to identify specific products or amounts beyond the paper receipt described above.

In section 3.2.2 we describe another prototype designed for another supermarket chain, where digital receipts can be accessed by logging into a portal and downloading a pdf file which can then be parsed in order to extract purchase data. This solution suffers from the problem of requiring manual handling and is therefore likely not something a customer will do for an extended period time. However, downloading and parsing several receipts does not take much longer time than one receipt, and internal test showed that downloading and parsing one year's worth of receipts takes about 20 minutes, thereby making it possible to analyze a long period of purchases. However, the most important advantage with this solution is also that these receipts are not just replicas of the paper receipt, but also contain EAN codes and PLU codes (Product Lookup Codes, see section 2.2.1) making it possible use product databases to extract much more information as described in sections 2.3.1 and 2.3.2.

A problem with both solutions these solutions for handling digital receipts is that each solution is closely tied to one specific supermarket chain, so a customer buying food at different supermarket chains would only get a partial image of their purchases. A better solution would be if the supermarket chains provided a service where the customer could authorize third party services to have direct access to that customer's purchase data. An outline for how such a system might work is described in [11], where issues of privacy and trust are discussed.

2.1.2 Extracting Data Without Receipts

One way to tackle the problem of solutions designed to a specific supermarket chain can be to bypass the receipts entirely and target the products themselves. We have developed two more prototypes where the customer can use a scanner or mobile phone to scan the barcodes on the products when unpacking the food at home, described in sections 3.1.2 and 3.2.1. These solutions suffer from the problem of requiring a substantial amount of manual work to extract the data, and also suffer from the problem that products without a fixed EAN/barcode cannot easily be added, but the solutions does not suffer from the problem of requiring different systems for different supermarket chains.

2.2 Identifying Products

After collecting data of what food you have bought, the next task is to identify these products and how precisely this can be done. We suggest a three-level taxonomy, where food can be identified on a "product type" level (i.e. "an apple"), on a "product level" (i.e. Kellog's cornflakes, 300 g") or on an "instance level" (i.e. "This particular piece of meat"). The more specific the identification is, the higher the possibility to accurately find data about the product is. Below, some current standards for identifying products are presented along with limitations on what kind of information is available with that method.

2.2.1 Product Type Level - PLU Codes

PLU codes or Product Lookup Codes is a global coding system used to uniquely identify bulk products sold in grocery stores or supermarkets. It is a four- or five-digit number associated with a specific type of (mostly) fruit or vegetable. These numbers can be keyed into point of sales systems by cashier and are sometimes available on receipts. Some number series are global and can therefore be easily identified, whereas other series are local and can be used by local retailers for specific products. Prefixing a four-digit number with "9" indicates an organic product.

While PLU codes makes it possible to discern between for example Granny Smith and Royal Gala apples, no individual differences within the PLU category can be made, for example country of origin, how the product was transported, when it was harvested or how it has been stored. This makes it difficult to get specific details about for example the total carbon footprint of the specific product, or local differences in nutritional content based on production methods, and more standard reference values must be used. Differences in carbon footprint can be substantial, as discussed in section 2.3.1 [16].

2.2.2 Product Level - GTIN and EAN Barcodes

GTIN numbers or Global Trade Item Numbers are 13-digit numbers used to uniquely identify products (but not instances of products). They are often found on food with a corresponding barcode (EAN code) and is used by food retailers for keeping track on sales and supply. Since GTIN numbers are global it is possible to both register and access various data about the products from different databases, such a nutritional data, information about origin, packaging, various markings and in some cases CO_2 emissions. However, the GTIN code is tied to the product level, not the instance of the products, making it impossible to know for example the origin of the beans in most varieties of ground coffee, which varies from day to day.

EAN codes are also used for food where the price is based on weight, and the weight can vary, for example for cheese. The price or weight is embedded in the EAN code. The number series used for these kinds of products are national, not global and are therefore not globally unique.

2.2.3 Product Instance Level - 2D Barcodes

The 13-digit limit of EAN and GTIN codes makes it impossible to supply additional information about the specific instance of the product. However, this will be addressed in the 2D barcodes that are being developed, and that will be available in 2027 [7]. This standard can add much more data, such as an URL, the expiry date or the batch number of a product, or even a unique identifier for each specific instance of each product, enabling far better possibilities for provenance, tracking and getting specific data about the instance of the product.

2.3 Finding Data About Identified Products

After the products have been identified as specifically as possible, the next stage is to find relevant data connected these products. Using the same taxonomy, we look at identifying food at the product type level, the product level and the product instance level.

2.3.1 Product Type Level

The product type level is when you have a generic type of product such as “an apple” or a meal such as “lasagna”. There are data and databases available with information about for example typical nutritional values and environmental impact on a product type level. For environmental data in a Swedish context, we have for example “Mat- klimatlistan” [18], a list developed by the Swedish University of Agricultural Sciences with typical CO_2e emissions for 41 common product types. The research institute RISE also has the open access “RISE Food climate database” [5] with 44 product types. Both these lists are adapted for typical Swedish conditions. In Denmark, there is an open data set called “The Big Climate Database” that includes data for climate footprints of 500 different food products broken down by stages of production .

There are also compilations on a global level such as a list included in the review by Poore and Nemecek [16], with data about land use, GHG emissions, acidifying emissions, eutrophying emissions, freshwater withdrawal and stress-weighted water use for 43 different product types. This dataset makes it clear that there are large variations even within a specific product type. For example, the listed GHG emission from the lowest decile of apples is 0.3 kg CO_2e/kg product but the highest decile is 1.0 kg CO_2e/kg product. For bovine meat the lowest decile is 14.4 kg CO_2e/kg product whereas the highest decile is 735.1 kg CO_2e/kg product. These differences are both due to different production methods, but also due to different system boundaries when the LCA of the product type was calculated.

A problem with both the lists is that they still include only a limited set of general products, and that categories such as “other vegetables” can include a wide variety of products with very different characteristics. Maintaining an updated and relatively complete list of all possible product types would be a daunting task. One way to tackle that problem could be to have a crowdsourced database with product type data. We have implemented such a database, LCAFDB (Life Cycle Analysis Food DataBase) [10] and used it in some projects described later.

For nutritional data, the Swedish food agency has developed The Swedish Food Composition Database [26] which provides information on the nutritional composition for more than 2000 foods and dishes. It lists 56 data points for each item in the database, including for example nutrients, energy content, amount of water and amount of expected waste (like banana peels or avocado seeds). These values are very useful for finding data about product types where no data is available on a product level, and can also be used by food manufacturers as default values for products where the manufacturers have not done a more detailed analysis.

2.3.2 Product Level

The product level is when you have for a specific product such as “Kellog’s Corn Flakes, 500 g”, often identified by a GTIN number and a barcode. This level can be associated with much more specific data than on the product type level. Basic nutritional data about products are required by EU legislation to be listed on the label of the product, specifically the amounts of fat, saturates, carbohydrate, sugars, protein and salt. In addition, more data may be added on a voluntary basis, specifically mono-unsaturates, polyunsaturates, polyols, starch, fibre and a number of vitamins or minerals “present in significant amounts”. [6]

While this makes it possible to extract some relevant information, the fact that it is voluntary to list much of the information opens up for problems. Whereas some products of a specific product type may have listed for example the amount of Vitamin D in the product, another product of the same product type may not have listed the amount of Vitamin D, even though, the second product might have equal or even higher content of Vitamin D. If a product contains only one product type, such as cheese, it is possible to use generic nutritional data for the nutrients not explicitly listed on the label, which can be found in for example the database from Livsmedelsverket mentioned above. This is also possible if the product can be mapped to a specific dish listed in the database, such as “Lasagna with ham”. However, there are no requirements for food manufacturers to provide explicit mapping between their products and product types, requiring such a mapping to be done either manually or by some algorithm. There are furthermore several products which does not map any entries in the databases, making it impossible to extract complete nutritional information for several products. An example of how this can cause problems is presented in section 3.2.1.

Furthermore, another problem with data on a product level is that several products use foods bought on the world market, where several important aspects such as producer, production method, country of origin may vary for one specific product. For example, one specific coffee product might sometimes contain beans grown in Ethiopia and at other times grown in Brazil. While this can cause variations in nutritional values, the problem is greater for environmental aspects since variations in environmental impact can differ much even within one specific product type as mentioned in the previous section.

The above paragraphs have presented problems with determining correct nutritional and environmental data on the product level. The next problem is for a system to get access to this data. There is no official global database containing these kinds of data. The data found on the labels of products are however often published on a number of web pages, for example by online supermarkets. This data can be scraped by systems in order to create databases. There are legal issues regarding the ownership of this data and how it can be connected to GTIN numbers, making the legality of such databases problematic. However, legal issues are outside the scope of this paper.

There are initiatives of crowd-sourced product databases. One is the Open Food Facts database [15] where users can contribute data about products, and which has an open API. To date, this database includes more than 2 million products. This database contains nutritional data found on labels as described above, but does not include environmental data.

In Sweden, the database Consupedia [3] made by the company with the same name includes more 200 000 products, and apart from nutritional data also includes environmental data where such is available, for example carbon footprint and freshwater use. This database has been used by us in several projects described in section 3.

2.3.3 Product Instance Level

Some of the problems identified on the product level could in theory be solved if the products could be identified on a product instance level rather than on a product level. For example, if each instance of a product could be identified and it would be possible to know exactly where, when and how the food components of the product were produced, and that would enable much higher accuracy of the data. However, to implement this completely on a global scale would require the possibility to track the product itself, to track its constituent parts and furthermore to have much more data about all steps in both the production and transport chain, essentially requiring a full LCA analysis for each product instance which is currently not realistic.

One way to track products on the instance level could be to use blockchain technology. One commercial example is the Foodtrace Traceability Platform [23] with which product can be tracked at all stages where a digital signal can be connected to the product instance, including an interface to the customer accessible by a QR code. A customer initialized Swedish project is “Sustainable and transparent purchases of cod using block chain technology” [24]by the municipality of Helsingborg in Sweden. The municipality will be able to track where the cod was caught and processed, and then follow the transport chain including location, temperature and humidity. The students and the elderly eating the fish will be able to access this information by reading a QR code.

2.4 Consumption and Waste Data

In the previous sections, we described how to identify what food you buy and how nutritional and environmental data can be collected by doing that. The next, and hardest problem is to know what you eat.

The simplest case would be a 1) one-person household that 2) buys all food in one supermarket chain where all data is available as described in section 3.2, 3) only eats what is bought in these supermarkets and never provides food to anyone else, and 4) wastes no food. In this case it can be assumed that over time everything that is bought is also consumed, making it possible to calculate both nutritional content and environmental impact over time.

However, that will probably be a very unusual case since 1) many households are not one-person households, 2) many at least occasionally buy food from different supermarket chains thereby requiring several systems to be integrated, 3) at least occasionally eat food at a restaurant or provides food for someone else, and 4) that most people waste at least some food.

2.4.1 Multi-Person Households

In a multi-person household, it is possible that several members of the household buys food at supermarkets separately. These purchases must then be combined in order to provide a full coverage of food bought in supermarkets. From a technical point of view this problem is easy to overcome, but it would also require all household members to agree to do this. That can be problematic from an ethical view since all household members might not want to share all private purchases with the rest of the household.

A perhaps bigger problem, at least from a technical point of view, is how to allocate the purchased food between the family members. If the interest is mainly in the environmental impact, this might not be necessary since the whole household could have a common goal and interest in this, and make collective decisions. If the interest of the household members is in healthy eating it is more problematic. Food purchases that might contains suitable amounts of nutrients for the household as a whole might hide imbalances within the household, where different household members might eat more or less of certain products and therefore different proportions of nutrients. This could be solved by detailed registrations of all ingredients in each meals, and detailed measurements of who eats how much of each meal. This could be feasible since it is not uncommon to track calories when following different diets, but would require a substantial amount of manual labor and would again raise ethical questions about privacy.

2.4.2 Food Bought in Restaurants

It is difficult today to track nutritions and environmental impact of food bought at restaurants. The problem could be solved if the restaurant calculated these data and provided the customer with the data, either by providing a QR code or by registering the purchase on a customer's profile, in the same way as supermarkets register purchases on different customers. In that case, the purchase at the restaurant could be treated as any purchase at a supermarket. Buffets however are more difficult to handle.

If many food purchases are difficult to register, such as if eating often at restaurants or if doing some purchases in supermarkets where purchases cannot be tracked, one approach can be to assume that people eat as many calories per day as they should according to health recommendations, and then normalize the purchases accordingly. For example, if the members of a household are supposed to eat 7500 calories per day, and the registered purchases total 2500 calories per day, then all nutritional and environmental data could be multiplied by a factor 3. An example of such a system is presented in section 3.2.2.

2.4.3 Food waste

Finally, food waste is a potential problem. In the ideal case that all food purchases can be tracked, the environmental impact will remain the same regardless of food waste, but the health related aspects will be affected since not all nutrients that are purchased are also eaten. Tracking food waste is a complex project outside of the scope of this paper, but there are recent innovative examples to improve this, such as "Svinnkollen" [25], an app that has been tested in school canteens. The components of different dishes served during a day are registered by the school. The students then use the app to take a picture of the meal served, and an AI identifies the meal and the components of the meal on the plate. After finishing the meal, the student takes a picture of the plate again, the AI identifies possible leftovers including the amounts and kinds of leftovers. The school can then calculate what the students have eaten, the amount of nutrients, the carbon impact of the food wasted and can also get feedback about what kinds of food the students like and don't like.

3 Prototypes Developed

In two research projects, both started in 2019, we have developed a number of prototype systems which tackle several problems included in this paper, and have often the hard way come to realize several of the problems mentioned in previous chapters. This is a brief account of some of these systems, and in which ways limits in data availability have hampered the systems, and in some cases been overcome.

3.1 Food Inventory and Carbon Footprints

Food inventory systems, where households can keep track of what food they have at home has been identified as desirable in order to reduce unnecessary purchases and therefore to reduce food waste [4, 9, 27]. Furthermore, ways to inform household of the carbon footprint of their food purchases is an important component in trying to support people in making more sustainable food choices. In both these cases, data about the food purchased is needed. Green Receipts and Green Cobra are two prototype systems we have developed addressing these issues.

3.1.1 Green Receipts

This system set out with the goal of identifying products purchased at a supermarket chain where receipts were sent out by email, and to calculate and give immediate feedback about carbon footprint of the purchases.

The main problem with this system was to identify which products were purchased. The only information on the receipts were the string found on the paper receipts, and in the case of food bought in bulk, also the weight. Some bulk products such as fresh fruit and vegetables could easily be identified and in these cases also the weight. These could be mapped to product type level as described in section 2.3.1, where data was found in the LCAFDB database [10]. Other products were more problematic.

One problem was that many product names were not easily identified by an algorithm, neither as products nor product types. For this purpose, a lookup database was included in the system, where the user could map unidentified products to product types. When the same string occurred later, the system would assume that the string was mapped to the same product type.

The other problem was that in some cases the product type could be identified, but not the weight. This was handled in a similar manner, where the user could enter a weight for the product. However, both these cases are not very robust since completely unconnected product types could have very similar names, and that the one string could be attached to a product type but where different products had different weights.

Nevertheless, the system was found to be useful on a general level, since the user could be made aware of both products with high and low carbon footprint per kg of product, and which products caused high and low impact based on the amount purchased, illustrating that availability of even very limited data can be useful [20]. The interface is shown in Figure 1.

Figure 1: Carbon footprint per kg is indicated by the color, total carbon footprint of the product indicated by g CO_2e

3.1.2 Green Cobra

GreenCobra was a food inventory system, where the user used a hand scanner at home to register food purchased when unpacking the food purchased and deregister the food when the product had been used [1]. The hand scanner could read the EAN code and when that code was connected to a GTIN number (i.e. the product level as described in 2.2.2), the product could be looked up in the Consupedia database, where various information could be found including carbon footprint. However, products without a barcode/GTIN such as fruit and vegetables could not be scanned, and products where the cost or weight were embedded in the barcode as explained in section 2.2.2 were not in the database.

Another problem was that although the database contained more than 200.000 products, only about 80% had data about CO2 footprint, making the data incomplete.

The final problem was about rights to use data in the database. While the Consupedia database included carbon footprint for most products, Consupedia was restricted in how that information could be exposed to their customers. Any use where CO_2e/kg of a product could be derived by a user was not allowed, which restricted our use to showing the quintiles of the product's carbon emissions. This meant the user could only see the footprint on a scale from 1-5, as shown in Figure 2. This also made it impossible to show the user the total carbon footprint of the user's last purchase.

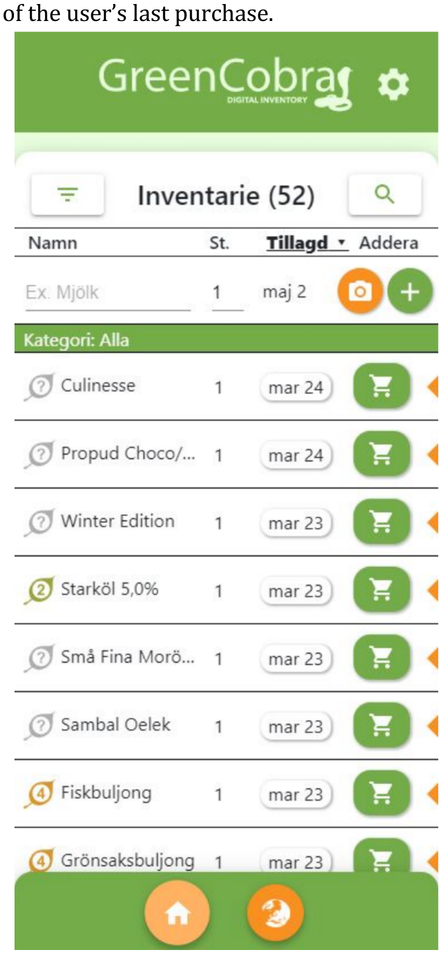

Figure 2: The GreenCobra food inventory system where the quintile of the carbon footprint per kg of product is shown in the leaves to the left.

3.2 Nutritional Data and Carbon Footprints

One result from these studies was that even people with deep environmental concerns chose what to eat not only based on environmental impact, but also on how healthy the food is and the price of the food. This led up to two more prototypes.

3.2.1 Foodprint

The Foodprint system was developed as an attempt to provide a simple interface for showing a user both the carbon footprint of a product and how healthy the product was. As a measure for health effect, we chose the NRF11.3 index [21], which is a nutrition index where 11 healthy nutrients and 3 unhealthy nutrients are combined to give a healthiness score. The software was based on GreenCobra above but with the difference that barcodes could be scanned with a mobile phone, and that the Consupedia database was also used to collect nutritional data about the products. The climate score calculated in GreenCobra was then combined with the NRF11.3 nutrition score to give a combined “goodness” result of the product.

A problem discovered was that some product types had registered the voluntary data for some vitamins, whereas other similar products had not entered this data, even though it is likely that they were similar. This resulted in that the products with more nutrients entered in the database were given higher score since any addition of these vitamins, however small, increases the NRF11.3 value. This resulted in an unfair disadvantage for products where this data had not been entered, which was discovered by the participants. A possible solution could be if the products were mapped to a more extensive nutritional database such as the one from Livsmedelsverket mentioned in section 2.3.1, and that if no data was found regarding a specific nutrient, the more general values found in the Livsmedelsverket database could be used. However, no such mapping exists, and would likely have to be done by the manufacturers or by crowdsourcing.

Even though these problems with the validity of some data existed, the system worked relatively well and provided the users with valuable insights [2]. One of the views is shown in Figure 3.

3.2.2My Foodprint

The final system presented is the My Foodprint system, that was developed to give recommendations of how to optimize food purchases so that nutritional requirements are met, and at the same time minimize either the cost or the carbon footprint. Data was collected from digital receipts which contained EAN codes for products. Since no API was provided the procedure to extract the data from the receipts was rather cumbersome. Pdf files of each receipt first had to be downloaded from an app, and then uploaded to the system where the pdf files were parsed and EAN codes extracted. Fruit and vegetables were registered with their PLU code, which made it possible to look up data on a product type level. The whole procedure took about 10-20 minutes for one year’s receipts.

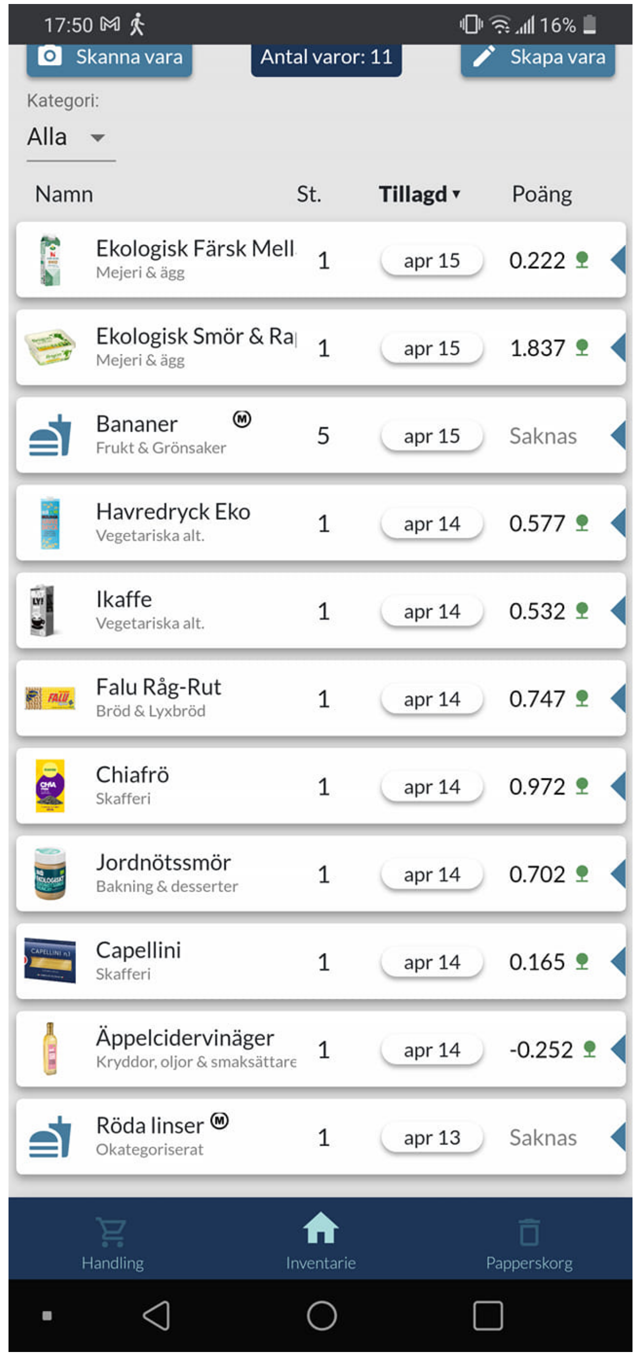

Figure 3: The Foodprint system, where a combined score for carbon footprint and nutritional value is presented in the right column.

Since the data did not contain all food purchases for a household, the purchases were normalized by assuming the purchases were typical and that household members ate the recommended number of calories each year.

After the data was in the system, the user would indicate how much they were willing to change their current purchasing habits, for example increase or decrease the amounts purchased of each product by a maximum of +/- 50%, and then enter whether they wanted to minimize carbon emissions or cost. This is a standard optimization problem that can be solved by linear programming.

We did however encounter a number of problems. Again, several, but not all, products did not contain data about for example vitamins. This resulted in that the system calculated the products to be less healthy than they actually were and resulted in that the optimization problem had no solutions. In the end we had to restrict the nutrients to the ones mandatory on product labels, that is saturated fat, sugar, protein and salt. Saturated fat, sugar and salt are the three nutrients that are considered unhealthy in the NRF11.3 nutrition index, and out of the 11 nutrients considered healthy, only protein could be used.

Furthermore, food with a barcode where price or weight was integrated in the barcode could not be used.

The system in its current form is partly useful, since it can give an indication about which products have high carbon footprints and at the same time contain unhealthy nutrients, but since it does not take into account many vital nutrients it cannot be used as a tool to suggest a full diet. The interface is shown in Figure 4.

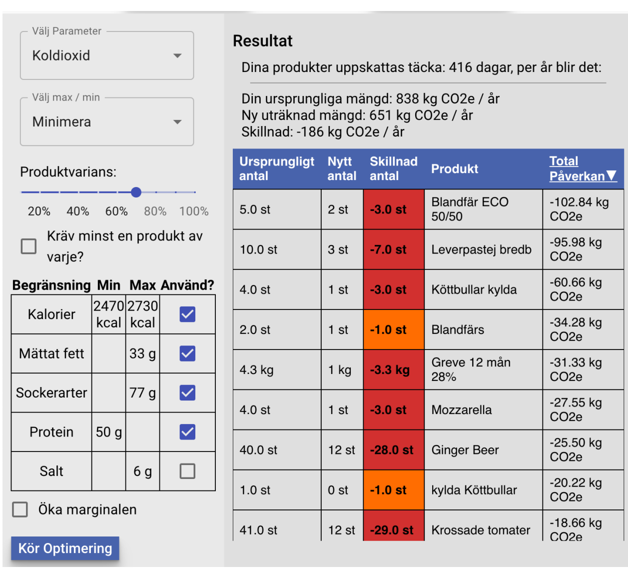

Figure 4: The My Footprint system, showing which ingredients would have the highest impact on CO2e emissions if they were decreased by a maximum 70%.

4 Discussion

In this paper, current technologies used to collect product and purchase data have been presented, along with a number of prototypes that could use data to support customers to live within the limits of a sustainable food system. Each system has encountered problems due to current limitations in data access and content. However, some of the problems in one system is overcome in another system, or are just a question of getting legal access to data or getting more complete data. So what could a system that combines all current solutions and where the solvable problems are solved be able to do, an "Omnifood" system to rule them all?

4.1 Access to Your Purchase Data

As shown in the My Foodprint project, it is not a technical problem to give customer access to what they buy. The supermarket has data on either the product type level or the product level, and also has access to weight and price of the products. The question is more related to whether the supermarket chains are willing to supply the customer with detailed data about their own consumption. Since this data is very valuable for the supermarket chains it might not be something they give up voluntarily. While GDPR legislation stipulates that customers should have access to the data companies have about the customer, this in practice has been very difficult to apply when it comes to consumption data.

With the upcoming 2D barcodes and the use of blockchain technology it will be possible to access much more data on the product instance level.

Ideally the purchase data should be accessible using an API and that customers can authorize different services to access their personal data as described in [11]. This would give the consumer the choice of determining which data is shared with which companies, and would allow data from several different sources, like several supermarket chains, to be combined in order to provide almost full coverage of purchases, except for small-scale purchases at for example markets.

4.2 Access to Product Data

Getting access to product data is not a technical problem, and could be handled with the solutions available. However, validity of the data can be problematic. The use of default values where no more specific data is available is reasonable, but which default values to use is not a simple question. For example, how to do LCA analyses of products are not standardized, and different choices of system boundaries often makes big differences, as will different production methods. A producer knowing that their product is worse than the default value could then choose not to provide their specific data, but instead use the default values making the product appearing better than it is.

There are also organizational and legal problems. The use of data connected to GTIN numbers is restricted and certain databases are not public, making it hard to build and maintain reliable and complete databases of products.

4.2.1 Access to What You Eat

Getting data about what you actually eat is the most difficult problem to overcome, since most options would require extensive amounts of manual registering. This makes it difficult to allocate environmental impact and nutritional data between household members, except in case of the one-person household.

4.3 Privacy and Ethical Issues

Finally, access to what people eat can be sensitive information. Employers might want to check out if potential employees have healthy eating habits before hiring staff, and insurance companies might not want to insure people with unhealthy lifestyles. Also, information about what we eat can be used for commercial purposes. There is also the question of privacy within a household. To have full coverage of what a household buys would require purchase data from several family members to be combined, which could be a problem if you purchase something you don't want other members to know that you have bought.

4.4 Possibilities of an Omniscient Food System

So, to wrap things up, what possibilities are there if a system would have access to all the data current and near-future technologies could offer about the food we buy and eat, and how that this connect to limits?

It would first of all be possible to get a relatively complete picture of the environmental footprint of the food we buy and eat, but with certain limits (see below). This environmental impact can be broken down into desired level of detail on product types, products or even product instances, and can provide valuable feedback to the engaged consumer wanting to minimize their environmental impact.

It could furthermore be possible to provide support for what to eat more and less of in order to stay within your personal financial limits, health limits and desired environmental limits, by using linear programming as described in section 3.2.2. Such a system could be valuable to everyone with external or self-imposed limits on food consumption.

However, no system, however technically advanced, can be better than its underlying models and assumptions. For example, the lack of standardized and well-established methods for how to do comparable life cycle analyses and how to quantify social sustainability makes it difficult to determine the validity of the environmental and social impact the food choices made. Furthermore, the food system is very complex, and even if there was such a thing as an agreed upon LCA, it could not cover all possible outcomes on all parts of the food system, where changes in the production methods of one product can have other impacts in completely different parts of the food system in another part of the world. Even with complete data about what we buy and eat, these underlying assumptions and models will provide researchers with many sleepless nights before an omniscient food system could be developed as the one solution the world has looked for that will solve all problems related to food consumption.

Finally, we have the question of how much interest there would be to actually use a system like this. In the area of HCI and energy, Strengers has described the “Resource man” as an image of how the energy sector looks upon their customers. “Resource Man is interested in his own energy data, understands it, and wants to use it to change the way he uses energy. He responds rationally to price signals and makes informed decisions based on up-to-date and detailed data provided about the costs, resource units (kilowatt hours), and impacts (greenhouse gas emissions) of his consumption. For these tasks he needs information…” [22]. Few energy consumers are like “resource man” and most likely few food consumers would be like a “food resource man”. Furthermore, such an omniscient system would make it painfully obvious to the consumer how much sensitive personal information about us is out there, and could make us feel like the year was 1984.

So while there are several potential uses of all the food data described in this paper, it is not necessarily the case that all these data should be used. There are after all limits both to what a computer system can solve, and what we would like a computer system to solve.

ACKNOWLEDGMENTS

This paper has been funded by the Swedish Energy Agency and Vinnova.

REFERENCES

[1] Bergqvist, H. et al. 2019. Inventariesystem för minskat matsvinn i hushåll En teknisk designstudie med teoretisk grund i beteendevetenskap En teknisk designstudie med teoretisk grund i beteendevetenskap. (2019).
[2] Bergqvist, H. 2021. Supporting consumers in making sustainable food purchases. KTH Royal Institute of Technology.
[3] Consupedia - A more transparent and sustainable world: 2022. https://consupedia.com/. Accessed: 2022-04-05.
[4] Farr-Wharton, G. et al. 2014. Food talks back: exploring the role of mobile applications in reducing domestic food wastage. Proceedings of the 26th Australian Computer-Human Interaction Conference on Designing Futures: the Future of Design (2014), 352–361.
[5] Florén, B. and Sjons, J. 2020. RISE Food climate database.
[6] Food Labelling Information System (FLIS): https://ec.europa.eu/food/safety/labelling-and-nutrition/food-labelling-information-system-flis_en. Accessed: 2022-04-05.
[7] GS1 2022. Global 2D Programme-Foundational FAQs Why should I switch from EAN/UPC (1D) barcodes to 2D barcodes?
[8] Har ungdomar näringsbrist? https://www.livsmedelsverket.se/matvanor-halsa--miljo/matvanor---undersokningar/riksmaten-ungdom/har-ungdomar-naringsbrist. Accessed: 2022-04-04.
[9] Hedin, B. et al. 2019. A systematic review of digital behaviour change interventions for more sustainable food consumption. Sustainability. 11, 9 (2019), 2638.
[10] Hedin, B. 2017. LCAFDB-A Crowdsourced Life Cycle Assessment Database for Food. The Fifth IFIP Conference on Sustainable Internet and ICT for Sustainability Funchal, Portugal--December 6-7, 2017 (2017).

[11] Isaksson, E. and Hedin, B. 2018. Smart Interactions for the Quantified Self. ICSLE 2018, The International Conference on Smart Learning Environments , March 18-20, 2018, Beijing, China (2018).

[12] Jäckle, A. et al. 2019. Participation in a mobile app survey to collect expenditure data as part of a large-scale probability household panel: Coverage and participation rates and biases. Survey Research Methods (2019), 23–44.

[13] Livsmedelsverket 2015. Protein - hur mycket är lagom.

[14] Obesity and overweight World Health Organization. https://www.who.int/news-room/fact-sheets/detail/obesity-and-overweight: 2020. .

[15] Open Food Facts - World: https://world.openfoodfacts.org/. Accessed: 2022-04-05.

[16] Poore, J. and Nemecek, T. 2018. Reducing food's environmental impacts through producers and consumers. Science. 360, 6392 (Jun. 2018), 987–992. DOI:https://doi.org/10.1126/science.aaq0216.

[17] Read, B. 2019. Respondent burden in a mobile app: Evidence from a shopping receipt scanning study. Survey Research Methods (2019), 45–71.

[18] Röös, E. 2014. Mat-klimat-listan.

[19] Sainz-De-Abajo, B. et al. 2020. FoodScan: Food Monitoring App by Scanning the Groceries Receipts. IEEE Access. 8, (2020), 227915–227924. DOI:https://doi.org/10.1109/ACCESS.2020.3046031.

[20] Shin, O. and Clemedson, E. 2019. Direktfeedback för att uppmuntra till hållbar matkonsumtion.

[21] Sonesson, U. et al. 2019. Dietary-dependent nutrient quality indexes as a complementary functional unit in LCA: A feasible option? Journal of Cleaner Production. 211, (2019), 620–627. DOI:https://doi.org/10.1016/j.jclepro.2018.11.171.

[22] Strengers, Y. 2014. Smart Energy in Everyday Life: Are You Designing for Resource Man? Interactions. 21, 4 (2014), 24–31. DOI:https://doi.org/10.1145/2621931.

[23] Supply Chain Traceability Software Platform: https://www.catsquared.com/foodtrace. Accessed: 2022-04-05.

[24] Sustainable and transparent purchases of cod using block chain technology - Innovation in Helsingborg: https://innovation.helsingborg.se/en/initiative/blockchain-for-savings-and-contract-follow-up/. Accessed: 2022-04-05.

[25] Svinnkollen 2020 - ForskarFredag: https://forskarfredag.se/forskarfredags-massexperiment/svinnkollen/. Accessed: 2022-04-05.

[26] The food database: 2021. https://www.livsmedelsverket.se/en/food-and-content/naringsamnen/livsmedelsdatabasen. Accessed: 2022-04-05.

[27] Woolley, E. et al. 2016. Manufacturing Resilience Via Inventory Management for Domestic Food Waste. Procedia CIRP. 40, (Jan. 2016), 372–377. DOI:https://doi.org/10.1016/J.PROCIR.2016.01.070.