**Horizon 2020**
**INFRADEV-1-2014 - Design studies**

**RICHFIELDS Working Package 5**
Deliverable D5.3

# Development of a quality evaluation framework for consumer generated food purchase data

**Date delivered:**
**M14**

**Author:**
Anne Normann, Susanne Ekman, Naomi Klepacz, Marcus Maringer, Anouk Geelen, Muriel Verain, Monique Raats

**Deliverable lead beneficiaries:**
SP Food and Bioscience


This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 654280.

| **Project** | |
|---|---|
| **Project acronym:** | RICHFIELDS |
| **Project full title:** | Research Infrastructure on Consumer Health and Food Intake for E-science with Linked Data Sharing |
| **Grant agreement no.:** | 654280 |
| **Project start date:** | 01.10.2015 |
| **Document:** | |
| **Title:** | Development of a quality evaluation framework for consumer generated food purchase data |
| **Deliverable No.:** | D 5.3 |
| **Authors:** | Anne Normann, Susanne Ekman, Naomi Klepacz, Marcus Maringer, Monique Raats, Anouk Geelen, Muriel Verain |
| **Reviewer:** | Karin Zimmermann – Project Coordinator<br>prof. dr. ir. Pieter van't Veer – Scientific Coordinator |
| **Start date:** | 1.10.2015 |
| **Delivery date:** | 01.12.2016 |
| **Due date of deliverable:** | 31.07.2016 |
| **Dissemination level:** | PU |
| **Status:** | |

| **Change history:** | | |
|---|---|---|
| Version | Notes | Date |
| Final | | 01.12.2016 |

**Karin Zimmermann**

**Project Coordinator**

**prof. dr. ir. Pieter van't Veer**

**Scientific Coordinator**



This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 654280.

## Summary

The overall aim of RICHFIELDS is to design a Research Infrastructure (RI) and data platform for the collection, integration, processing and sharing of consumer generated data related to food intake activities. In order for the data to be valuable to users of RICHFIELDS it is essential that factors influencing the quality of this data are identified and thereby visualize the potential opportunities, as well as the gaps and needs, with the data as part of the collection, integration and dissemination process.

A set of quality criteria was formulated for the evaluation and inventory framework of the consumer generated food intake activities, within the areas of scientific relevance and technical and legal governance. Furthermore, the result of this deliverable should also provide structure and guidance for the data collection and inventory of consumer generated food purchase tools (task 5.1).

A literature search has been conducted and existing quality frameworks of eHealth and mHealth applications have been summarized in order to create the quality framework. Quality criteria from that overview were selected based on the significance for the quality dimensions, data management and legal governance. To evaluate the relevance of the selection of quality criteria, experts in the relevant fields of Law and ICT were contacted. Based on the experts' opinions the selection of quality criteria was adjusted. The work also continued parallel to the actual inventory (task 5.1), adding variables/inputs to the criteria alongside increased knowledge about different tool types and what consumer generated purchase data they potentially generated. However, existing quality frameworks are rather general in nature with respect to scientific relevance and do not focus on specific scientific fields such as those relevant to RICHFIELDS. Thus, it was necessary for the assessment of quality within RICHFIELDS to create a unique set of criteria. The selected quality criteria are thought to be relevant and comprehensive across the needs and requirements of the various disciplines involved in designing the blueprint of the RI and data platform.


This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 654280.

# Contents




This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 654280.

# 1. Introduction

## 1.1 BACKGROUND

In light of digitalization, knowledge generation is changing all spheres of activity including social interactions, business and health care resulting in what is often referred to as "Big Data". Such data has been characterized by the four dimensions of volume (the amount of data being generated), velocity (the speed at which data needs to be processed), variety (issues relating to data heterogeneity), and veracity (the degree of uncertainty in the content of user generated data) (Schroeck et al., 2012). The Horizon2020 project RICHFIELDS recognizes that the open data movement in research and innovative ways of data collection including user-generated (big) data provide unprecedented possibilities to study diet, lifestyle, health and their determinants. Data can be collected by using new media, e.g. in the social space (the Web, GIS) and real-time (apps, wearables, GIS, sensors) at the individual and group (e.g. household) level. These data could provide valuable information on the association between determinants and dietary intake which is of high societal and scientific relevance.

Particularly in the area of infectious disease monitoring, the use of user-generated data has been heralded as an opportunity to improve public health surveillance (Velasco et al 2014). Health agencies have been reluctant to incorporate these data sources into their systems because many technical issues have not yet been addressed (Velasco et al 2014). Considerations of data protection and privacy, such as legal and ethical implications related to using Internet and social media data are also needed (Velasco et al 2014). In the area of diet and health, researchers have recently used data collected through Twitter (Abbar, Mejova and Weber, 2014; Fried et al., 2014) and Instagram (Mejova, et al, 2015; Sharma and De Choudhury, 2015) to study food consumption patterns. Weber and Achananuparp (2016) used data from public food diaries collected using the application MyFitnessPal to construct models to predict whether users will or will not meet their daily caloric goals.

Food choice operates at physical, biological, psychological, and sociocultural levels (Sobal, 1991), all which operate simultaneously and interact (Sobal, Bisogni and Jastran, 2014). In scientific research, data is collected in controlled conditions to provide insights. The types and sequences of food-related behaviours are depicted in Figure 1 and include the acquisition, preparation, serving, eating, storage, giving away of and cleaning up of food.

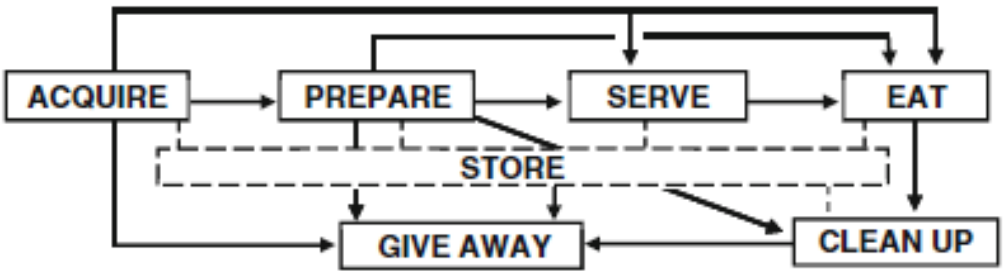



Figure 1. Summary of types and sequences of food behaviours (from Sobal and Bisogni, 2009)


This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 654280.

Phase 1 of the RICHFIELDS project seeks to identify food-related data that is being actively or passively generated by consumers through the use of tools such as apps and sensors. "Outside the research environment", people are generating data through everyday food-related activities. These might include banking transactions from which food-related purchase can be estimated, food related (e.g. recipes, restaurant reviews) search behaviour on the internet and the use of apps to record food intake or disclose food-related images or text. The large-scale generation of such data could have the potential to provide data for the purpose of research which can provide insights regarding food choices and can relate to the purchase, preparation and consumption of food.

## 1.2 AIM

The aim of this deliverable (D5.3) is to formulate a set of quality criteria for the evaluation of this consumer generated food data in terms of its scientific relevance and technical and legal governance. These three areas were selected as indicators of quality as they allow for the assessment of data in relation to key questions relating to food purchase, preparation and consumption/lifestyle behaviour (i.e., What/Who/Why/How and Where), in addition to assessing the legal limitations, organizational restrictions, confidentiality and privacy concerns related to collection, integration and dissemination of consumer generated data and the technical protocols and standards for data access and data processing. Information about these topics is crucial for developing the blueprint of the data platform and its data governance structure.

In addition to providing a framework for the evaluation of the data, the result of this deliverable should also provide structure and guidance for the data collection process of deliverable 5.1, which is an inventory of consumer generated purchase data tools (see chapter 1.3 below). More specifically, the framework will provide operationalisations for each quality criteria in the form of a set of relevant questions that should be answered for each tool included in the inventory of deliverable 5.1.

The aim of this deliverable is not to create an exhaustive list of criteria for the validation of 'Big data' sets for their potential use in social science research. Although, it is acknowledged that such validation is of course crucial for the use of 'Big Data' in social science research and warrants closer examination in relation to specific research questions. Instead, the quality criteria as set out in this document aim to access whether the data produced through individual tools (i.e., smartphone apps, websites, and sensors) has the potential for use in consumer research. That is to say, is the consumer generated data 'fit for purpose'?

Furthermore, we believe that such a 'tool' or source level validation of quality should form an integral part of the Research Infrastructure and should therefore be considered and incorporated into the overall design of the data platform.


This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 654280.

### 1.3 RICHFIELDS INVENTORY MANAGEMENT SYSTEM (RIMS)

The RICHFIELDS Inventory Management System (RIMS) was created in response to Task 5.1 which required the creation of an inventory of types of purchase, preparation, consumption and lifestyle data, and data collection methodologies. In brief, RIMS is an online management system for the storage and assessment of tools that produce consumer generated food and/or beverage purchase [*and preparation and consumption*] data that could potentially be of use to researchers. RIMS comprises two component parts; [1.] a typology and categorisation of the tools stored within the inventory, and [2.] a list of quality criteria against which each tool can be validated. These component parts will now be described in sections 2 and 3 of this deliverable.

# 2. The typology

### 2.1 DESCRIPTION OF TYPOLOGY

The typology is a scheduled framework categorizing the tools in different groups. The typology at Level 1 is based on the definition of purchase (see chapter 2.2) and the sub-groups (Level 2) describe the purpose of the data collected, what recordable activities that are captured and the potential consumer data it generates (Level 3), Figure 2.

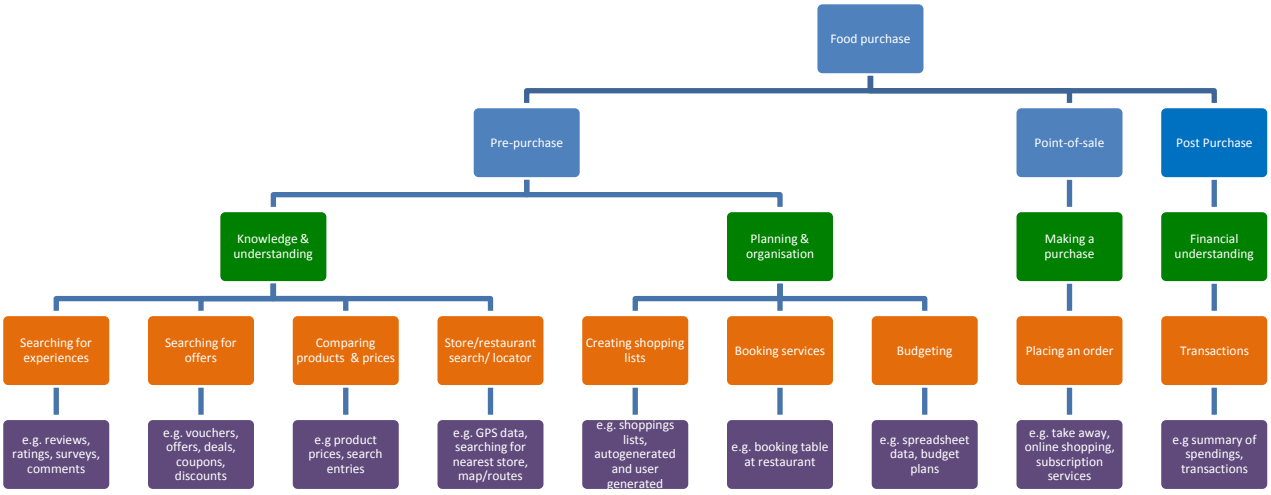



Figure 2. Typology of purchase

### 2.2 DEFINITION OF FOOD PURCHASE

The intention of Richfields Phase 1 is to cover the area of food intake activities. By this, WP5, WP6 and WP7 cover the activities purchase, preparation as well as consumption. WP 5 only covers purchase. It is of course well known that not all food we eat is purchased, it might be home gardening or meals at school, but food intake includes purchasing behaviour. The


This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 654280.

term “purchasing” has historically first and foremost been defined as the decision at the actual point-of-sale. However, purchase behaviour can instead be defined as a process that goes beyond the act of purchase at the product shelf. Rather, it includes different factors which can influence the consumer before, during or after a purchase decision (Solomon et al. 2013). Levy, Weitz and Grewal (2014, p.91) defined the process in 5 steps. It begins with the pre-purchase phase, which includes the recognition of a need/motive, a more or less intensive information search determined by the current type of buying decision, and an evaluation of different options (Howard & Sheth 1969, p. 25f). The pre-purchase phase includes processes where consumers compare prices-, groceries-, product-, service- and store related information, plan and decide what to buy or cook (Saarijärvi et al., 2014).

After the pre-purchase phase, the purchase decision at the point-of-sale is made. Finally, the consumer evaluates the buying decision in the post-purchase phase (e.g. Levy et al. 2014, p. 91). Regarding purchase within the scope of WP5 we primarily focus on the phases of pre-purchase, as well as the actual point-of-sale, because the post-purchase phase is partly covered in consumption (WP7), but also because it is assumed that post-purchase is the base for another (pre-)purchase phase (internal search for information due to memories and experiences).

## 3. Quality criteria

WP7 took the lead in the work of developing a framework to the quality criteria currently in RIMS. In order to create the quality framework a literature search was conducted. Private as well as public companies and institutions offer guidelines, services and infrastructures for reviewing, evaluating and certifying health applications, and the literature search was conducted on these existing quality frameworks. Quality criteria from that overview were selected based on the significance for the quality dimensions related to scientific relevance and legal and technical governance. In order to evaluate the relevance of our selection of quality criteria which reside outside of our own field of expertise (legal and technical governance), we contacted experts in the relevant fields of Law[1] and ICT[2] (one distinguished expert for each field of expertise). Based on the experts’ opinions the selection of quality criteria was adjusted. The work also continued parallel to the actual inventory, adding variables/inputs to the criteria alongside increased knowledge about different tool types and what data they potentially generate. However, existing quality frameworks are rather general in nature with respect to scientific relevance and do not focus on specific scientific fields such as those relevant to RICHFIELDS. Thus, it is necessary for the assessment of quality within RICHFIELDS to create a unique set of criteria.

[1] Prof. dr. Indira Carr, Faculty of Business, Economics and Law, University of Surrey, Guilford, Surrey, United Kingdom.

[2] Dr. Barbara Korousic Seljak, Computer Systems Department, Jožef Stefan Institute, Ljubljana, Slovenia.


This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 654280.

## 3.1 DEFINITION OF QUALITY CRITERIA

As described in RICHFIELDS deliverable D4.1, the assessment of data quality is a necessary stage enabling the extraction, cleaning and transfer of data into information. There are many guidelines available, on which principles of data quality should be addressed in the process of data extraction/translation into information and/or knowledge. RICHFIELDS will align all data information systems with a flexible approach to FAIR principles, so that ultimately all the data that flows in and out of RICHFIELDS is; Findable, Accessible, Inter-operable and Re-usable (e.g., Wilkinson et al., 2016). See more in D4.1; chapter 4.3.2 about data quality. We have also tried to use the FAIR principles to reflect on purchase data. Several criteria have for example been identified as relevant for assessing the quality related to the technical governance of the consumer generated food purchase data, see 3.1.3 below.

### 3.1.1 DESCRIPTIVE QUALITY CRITERIA

In order to effectively evaluate the quality of a tool, it is essential that certain key characteristics are identified, described and recorded within RIMS. These descriptive characteristics focus on the identification of the source of the information and classification of each tool according to the typology set out in chapter 2. The Descriptive Quality Criteria are found in table 1.

Each tool needs to be 'traceable', both in terms of the search strategy used to locate the tool and in terms of where additional information can be found about the tool (e.g., a website), or – in the case of apps – downloaded. Therefore, in addition to basic information, such as the name of the tool as it appears in iTunes, for each tool the 'search type' used to find the tool needs to be recorded, in addition to information about the 'search engine', 'reference tool' and 'search term'. Descriptions of these criteria and variables currently identified within RIMS can be seen in the top row of table 1. Further identifying factors about the tool are also collected. This is to assist in maintaining reliability, as it enables users to ensure they are using the same tool obtained via the same source. These additional identifying factors include; a copy of the 'tool logo', the 'tool description' as written by the tool developer, 'languages' supported by the tool as well as 'supported platforms'. In addition, information on how to access the tool is also recorded, including where the tool can be accessed or downloaded, the website and homepage for the tool and the name of the company or app developer who currently owns the tool. Descriptions of these characteristics can be found in the bottom row of table 1.

Information is also collected on the tool type. That is, whether the tool is an app, website or software. In the case of apps obtainable via the iTunes store, further information concerning the characteristics of these apps can be obtained directly via the apps' unique iTunes ID number. Thus, for each app – where possible – the iTunes ID number was recorded.

An important indicator for the quality of a tool is whether it meets the overarching aim of the RICHFIELDS project. That is to say, the tool must collect consumer generated data on, in this instance, food purchase. Thus the tool – and the data it collects - must be classifiable according to the typology set out in chapter 2. The descriptive quality criteria therefore list criteria for 'data types,' that is whether or not the app collects consumer generated 'food purchase data'. Further descriptive quality criteria allow for the identification of the consumers' 'goals' and also their 'behaviour' according to the typology.


This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 654280.

**Table 1. Descriptive criteria – how did we find the tools and what kind of tool is it?**

| **Criteria** *(sub-criteria)* | **Descriptive Question and/or RIMS user instructions** | **Criteria Description** | **Variables** |
|---|---|---|---|
| Search Type | How did you find the tool? | To identify the type of search strategy that was used to identify the tool. | Search Engine<br>Reference |
| *Search Engine* | *What search engine has been used to find the tool?* | *Variable contingent upon the select of the variable 'search engine'. Name of the specific search engine that has been used to identify the tool.* | *Appcrawl*<br>*Vinoza*<br>*Google Play*<br>*Google Search*<br>*iTunes*<br>*PubMed*<br>*Fnd.io* |
| *Search Term* | *What is the search you used to find the tool?* | *Variable contingent upon the selection of the variable 'search engine'. The search term or string of terms used to identify the tool.* | *Text Entry* |
| *Tool Reference* | *Where did you find the tool* | *Variable contingent upon the selection of the variable 'reference'. The scientific reference used to identify the tool (e.g., journal article).* | *Text entry* |
| *Tool Type* | *What category does the tool belong to?* | To identify the type of tool and the category to which it belongs. | App<br>Website<br>Sensor |
| *Query ITunes Store* | *If selected, after you save the item application information will be collected from ITunes search AP! And automatically inserted.* | *The iTunes ID number for the tool.* | *Text entry* |
| *Name* | *What is the name of the tool?* | *The name of the tool.* | *Text Entry* |


This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 654280.

| | | | |
|---|---|---|---|
| *Data Types* | *What type of data does the tool collect?* | To identify the type of consumer generate data that the tool collects. | Lifestyle<br>Purchase<br>Preparation<br>Consumption |
| *Category purchase* | *What purchase category does the tool belong to?* | *To identify the category relating to the type of consumer generated purchase data that the tool collects.* | *Knowledge & understanding*<br>*Planning & organisation*<br>*Making a purchase*<br>*Financial understanding* |
| *Subcategory purchase* | *What purchase subcategory does the tool belong to?* | *To identify the subcategory relating to the type of consumer generated purchase data that the tool collects.* | *Store/restaurant, search/locator (Knowledge and understanding)*<br>*Comparing products & prices (Knowledge and understanding)*<br>*Searching for experiences (Knowledge and understanding)*<br>*Searching for offers (Knowledge and understanding)*<br>*Booking services (Planning & organisation)*<br>*Creating shopping lists (Planning & organisation)*<br>*Budgeting (Planning & organisation)*<br>*Placing and order (Making a purchase)*<br>*Transactions (Financial understanding)* |
| Tool Logo | Add a URL to a logo image | The Uniform Resource Locator (URL) web 'address' for the logo image associated with the tool | Text entry |
| Tool description | Tool description | The description of the tool provided by the app developer and visible to the app user in iTunes. | Text entry |
| Languages | Which languages are supported? | A list of languages supported by the tool | Text entry |
| Download / Access URL | Where can the tool be downloaded or accessed? | The URL web address from which the tool can be accessed or downloaded. | Text entry |
| Company Name | What is the name of the company who owns the | The name of the company who owns the tool. | Text entry |


This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 654280.

| | | | |
|---|---|---|---|
| | tool? | | |
| Website URL | What is the web address of the tool? | The URL web address for the tool. This website may be a supporting website or a website for the company or web developer who owns the tool. | Text entry |
| Has home page | Does the tool provide a link to a working home page? | To identify whether or not the tool has a working home page | Yes<br>No |
| Supported platforms | What are the supported platforms? | To identify which platforms are supported by the tool. | Kindle<br>Watch OS<br>Android wear<br>Blackberry<br>HTML5<br>Windows Phone<br>OSX<br>Windows<br>Android<br>IOS |

### 3.1.2 SCIENTIFIC QUALITY CRITERIA

Food choice behaviour in general is a seemingly simple, but in fact very complicated behaviour that is influenced by many interacting factors. Moreover, these factors each belong to the traditional domains of one of a large diversity of scientific disciplines and as a result each of these disciplines claims to have at least a partial answer to the central question in food choice research: ''*Why* does *who* eat *what*, *when*, and *where*?" (Köster, 2009). The complexity also goes for food purchase behaviour.

Some generic scientific criteria were created which further on was detailed into individual units for purchase, preparation and consumption respectively.
When recording a "Yes" for e.g. *Lifestyle Data, Situational Characteristics* and *Product characteristics* (table 2) the variables presented in table 3, 4 and 5 were displayed in RIMS respectively.

**Table 2. Generic scientific quality criteria**

| **Criteria** *(sub-criteria)* | **Scientific Question** | **Criteria Description** | **Variables** |
|---|---|---|---|
| Lifestyle Data | What type of lifestyle data does the tool collect? | To identify the type of consumer generated lifestyle data collected by an app relating to purchase, preparation and/or consumption behaviour. | Text Entry |


This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 654280.

| | | | |
|---|---|---|---|
| Situational Characteristics | Does the tool collect information about the situation of a consumer? | To identify whether or not the app collects data regarding the context and/or situation in which the consumer generated data collected by the app was generated. | Yes<br>No |
| *Type of situational characteristics* | *What type of situational characteristics does the tool collect?* | *The type of context and/or situational data collected by the app.* | *Text Entry* |
| Product Characteristics | Does the tool collect information about the characteristics of the product what has been consumed, prepared or purchased? | To identify whether or not the app collects consumer generated data regarding the type of product purchase, prepared or consumed by the user. | Yes<br>No |
| *Type of product characteristics* | *What type of product characteristics does the tool collect?* | *The type of product characteristic(s) collected by the apps.* | *Text Entry* |
| External devices of same vendor | Does the tool support external devices owned by the vendor of the tool? | To identify whether or not the app supports external devices manufactured by the same company as the current tool with the express intent of being used in partnership. | Yes<br>No |
| *Device Type* | *What type of external devices does the tool support?* | *The type of external device(s) supported by the app.* | *Text Entry* |
| Data integration with partner tools | *Does the tool integrate data from other tools?* | To identify whether or not the app integrates consumer or non-consumer generated data (such as demographic data) from other tools included in RIMS. | Yes<br>No |
| *Partner tools* | *Of which other tools does the tool integrate data?* | *The name of the tool included in RIMS from which the current tool takes its information.* | *Text Entry* |

Criteria for life style and situational characteristics were included since self-monitoring/tracking generated data can be interesting from a consumer behavioural perspective. There is also a desire to use social media, to like, evaluate and/or share various aspects of food purchase activities. The concept of 'self-tracking' has recently begun to emerge in discussions of ways in which people can monitor and record specific features of their lives. Self-tracking is also referred to as lifelogging, personal analytics and personal informatics (Lupton, 2016).

**Table 3. Lifestyle quality criteria – what information is collected about the products that has been purchased?**


This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 654280.

| **Criteria** *(sub-criteria)* | **Variables** | **Variable Description** |
|---|---|---|
| *Lifestyle characteristics* | *GPS data* | *GPS data is data derived from the global positioning system and can include route, distance, and pace in once exercise and movement* |
| | *Location preferences* | *Some locations that is liked or wanted more than another location such as restaurants or bars* |
| | *Food preferences* | *Some foods that is liked or wanted more than another thing* |
| | *Comments* | *A comment is a statement of fact or opinion, especially a remark that expresses a personal reaction or attitude* |
| | *Posts* | *Shared messages in an online forum, newsgroup or connected friends* |
| | *Evaluations* | *Acts or instances of evaluating or appraising* |
| | *Notes* | *A brief record, comment or explanation written down to aid memory* |

**Table 4. Situational quality criteria – what information is collected about the products that has been purchased?**

| **Criteria** *(sub-criteria)* | **Variables** | **Variable Description** |
|---|---|---|
| *Situational characteristics* | *Venue name* | *The name of the venue a person is or was located* |
| | *Notifications* | *Acts or instances of notifying, making known, or giving notice; notice* |
| | *Connected users* | *Users which follow each other's progress, posts, comments or other sorts of shared information* |

**Table 5. Product characteristics criteria – what information is collected about the products that has been purchased?**

| **Criteria** *(sub-criteria)* | **Variables** | **Variable Description** |
|---|---|---|
| *Product characteristics* | *Food group* | *A collection of foods that share similar nutritional properties or biological classifications* https://en.wikipedia.org/wiki/Food_group *such as dairy products or bakery.* |
| | *Food* | *The type of food which has been purchased such as bread, cake, or steak* |
| | *Product* | *The specific food product which has been purchased such as Dunkin Donuts, Coca Cola or a Quaker Oats cereal* |
| | *Cuisine* | *A style or method of cooking, especially as characteristic of a particular country, region, or establishment* http://www.oxforddictionaries.com/definition/english/cuisine |
| | *Dish* | *A dish is a specific food preparation, a distinct article or variety of food, with cooking finished, and ready to eat, or be served* |
| | *Beverage* | *Any one of various liquids for drinking, usually excluding water* |
| | *Ingredients* | *Edible substances that are used in a dish or a product* |
| | *Energy content* | *Energy declaration on a food product or the energy content in a dish in kcal or KJ* |
| | *Nutrient content* | *Nutrition declaration contains as minimum the information of: energy value (in both kilojoules (kJ) and kilocalories (kcal)); and the amounts (in grams (g)) of fat, saturates, carbohydrate, sugars, protein and salt. EU regulation (EU) No. 1169/2011* |
| | | *Info for diverse human eating. Diet might alter for many reasons, e.g. for physical or mental* |


This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 654280.

| | |
|---|---|
| *Special diet* | *health concerns, religious or ethical reasons, weight loss, diseases etc.* http://medical-dictionary.thefreedictionary.com/Diets |
| *Allergen information* | *Any prepacked food or drink sold in the UK must clearly state on the label if it contains the following ingredients found on:* https://www.food.gov.uk/science/allergy-intolerance |
| *Cooking advice and instructions* | *The manufacturer's instructions for preparing the food. Instructions for use on a dry, fridge or a freezer product.*http://www.food.gov.uk/sites/default/files/multimedia/pdfs/gceone_class.pdf |
| *Storage condition* | *In cases where foods require special storage conditions and/or conditions of use, those conditions shall be indicated.* http://eur-lex.europa.eu/legal content/EN/TXT/PDF/?uri=CELEX:32011R1169&rid=1 |
| *Product origin* | *The place of manufacture, production, or growth where an article or product comes from.* http://www.thefreedictionary.com/origin |
| *Product volume* | *The amount of a product that is enclosed within a container e.g. litre, centilitre or millilitre.* http://www.oxforddictionaries.com/definition/english/volume |
| *Product weight* | *The amount or quantity a product weighs.* http://www.dictionary.com/browse/weight |
| *Visual properties* | *A sensory attribute of the food.* http://www.thefreedictionary.com/visual+properties |
| *Price* | *The amount of money for which something is sold.* http://dictionary.cambridge.org/dictionary/english/price |
| *Availability* | *Food availability is defined as sufficient quantities of food of appropriate quality, supplied through domestic production or imports, including food aids* |
| *No Information* | *Not enough information to determine the what unit method* |

In relation to the earlier mentioned questions about ''*Why* does *who* eat *what*, *when*, and *where*?" (Köster, 2009), chosen criteria had to be clarified regarding *what* m-services? in particular provide for consumers and *what* potential data it generates. M-services can support customers both during pre-purchase, in-store purchase and post-purchase processes; to decide what to buy, ease comparison of products, service- and store-related information, finding the right products, learn about nutritional information, store locator and manage shopping lists, self-scanning, pay for purchase at the checkout etc. (Saarijärvi et al, 2014). According to Shankar et al (2010); create shopping lists, search, compare, purchase and post-purchase activities were identified as mobile consumer attitudes and behaviours. Local search, shopping search, visual/camera phone search, voice search, mapping are added as activities by Kroski (2008), and of course also make an actual purchase. Table 6 visualizes the variable descriptions and explanations of inputs of the "what methods", "what consumer unit" "how much methods", "when", "where" and "how method" for the purchase specific criteria presented in table 8.


This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 654280.

**Table 6. Purchase specific criteria – What, where, how and when was it purchased?**

| **Criteria** *(sub-criteria)* | **Scientific Question** | **Criteria Description** | **Variables** |
|---|---|---|---|
| What was purchased? | Does the tool collect data on what was purchased? | To identify whether or not the app collects consumer generated data on 'what' domestic food(s) and/or beverage(s) have been purchased. That is, the app collects observable data relating to the nature and characteristics of purchased domestic food(s) and/or beverages. | Yes<br>No<br>No information |
| *Methods: What was purchased?* | *Which method(s) have been used to collect data on what was purchased?* | *To identify the method(s) of data collection used by the app to capture data relating to nature and characteristics of food(s) and/or beverage(s) purchased.* | *No information*<br>*Voice input*<br>*Order confirmation*<br>*Loyalty card*<br>*Manual input*<br>*Manual search*<br>*Scanned receipt*<br>*Financial transaction*<br>*Purchase history*<br>*Barcode scanning*<br>*Spectroscopic analysis* |
| *What was purchased: act or Intention?* | *Does the data about what was purchased refer to intentions to purchase and/or actual acts of purchase?* | *To identify whether the data on what food(s) and/or beverage(s) purchased captured by the app relates to an 'actual' act(s) that have taken place – or, whether the data relates to an 'intended' act(s). That is, data captured relates to an act(s) that have yet to take place.* | *Act*<br>*Intention*<br>*Both* |
| *What unit: purchase* | *What is the unit of measurement?* | *The unit in which the nature and characteristics of the food(s) and/or beverage(s) that have been purchased have been measured.* | *No information*<br>*Food box*<br>*Food*<br>*Food group*<br>*Product*<br>*Ingredients*<br>*Dish*<br>*Beverage*<br>*Money* |


This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 654280.

| | | | |
|---|---|---|---|
| *What is the consumer unit: purchase?* | *What is the consumer unit for which purchase have been measured?* | *The population for which data the nature and characteristics of the food(s) and/or beverage(s) that have been purchased have been captured.* | *No Information*<br>*Group*<br>*Household*<br>*Individual* |
| How much was purchased? | Does the tool collect data on how much was purchased? | To identify whether or not an app collects observable data relating to the quantity of domestic food(s) and/or beverage(s) that have been purchased. | Yes<br>No<br>No Information |
| *Methods: How much was purchased?* | *Which method(s) have been used to collect data on how much was purchased?* | *To identify the method(s) of data collection used by the app to capture data relating to the quantity of food(s) and/or beverage(s) purchased.* | *No information*<br>*Serving size*<br>*Portion size*<br>*Product volume*<br>*Product weight*<br>*Spending* |
| *How much was purchased: act or intention?* | *Does the data about how much was purchased refer to intentions to purchase and/or actual acts of purchase?* | *To identify whether the data on the quantity /amount of food(s) and/or beverage(s) purchased captured by the app relates to an 'actual' act(s) that have taken place – or, whether the data relates to an 'intended' act(s). That is, data captured relates to an act(s) that have yet to take place.* | *Act*<br>*Intention*<br>*Both* |
| When was it purchased? | Does the tool collect data on when the purchase took place? | To identify whether or not an app collects observable data relating to the time at which food purchase took place. | Yes<br>No<br>No Information |
| *Time unit: Purchase* | *In which unit(s) of time has purchase been measured?* | *The unit in which the time that purchase took place has been measured.* | *No Information*<br>*Date*<br>*Weeks*<br>*Exact time*<br>*Day Periods*<br>*Days*<br>*Months*<br>*Years* |
| *When was it purchased: Act or Intention?* | *Does the data about when food was purchased refer to intentions* | *To identify whether the data on when food(s) and/or beverage(s) purchased took place captured by the app* | *Act* |


This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 654280.

| | | | |
|---|---|---|---|
| | *to purchase and/or actual acts of purchase?* | *relates to an 'actual' act(s) that have taken place – or, whether the data relates to an 'intended' act(s). That is, data captured relates to an act(s) that have yet to take place.* | *Intention*<br><br>*Both* |
| Where was it purchased? | Does the tool collect data on where the purchase took place? | To identify whether or not an app collects observable data relating to the physical location in which the purchase of food(s) and/or beverage(s) took place. | Yes<br><br>No<br><br>No Information |
| *Location unit: purchase* | *In what unit has the location of purchase been measured?* | *The unit in which the physical location in which food purchase took place was measured/ recorded.* | *No Information*<br><br>*Venue Name*<br><br>*Geo - Coordinates* |
| *Where was it purchased: Act or intention?* | *Does the data about where was purchased refer to intentions to purchased and/or actual acts of purchase?* | *To identify whether the data regarding the physical location in which food(s) and/or beverage(s) purchased took place relates to an 'actual' act(s) that have taken place – or, whether the data relates to an 'intended' act(s). That is, data captured relates to an act(s) that have yet to take place.* | *Act*<br><br>*Intention*<br><br>*Both* |
| How was it purchased? | Does the tool collect information about how the purchase took place? | To identify whether or not an app collects observable data relating to how food(s) and/or beverage(s) were purchased. That is, the apps capture data relating to the mechanism of food purchase. | Yes<br><br>No<br><br>No information |
| *Method: How was it purchased?* | *What data does the tool collect about how the food was purchased?* | *To identify the method(s) of data collection used by the app to capture data relating to the mechanism of food(s) and/or beverage(s) purchase.* | *No information*<br><br>*In-store*<br><br>*Online shopping* |
| *How was it purchased: Act or intention?* | *Does the data about how was purchased refer to intentions to purchased and/or actual acts of purchase?* | *To identify whether the data regarding the mechanism of food purchase relates to an 'actual' act(s) that have taken place – or, whether the data relates to an 'intended' act(s). That is, data captured relates to an act(s) that have yet to take place.* | *Act*<br><br>*Intention*<br><br>*Both* |
| Occasion | Does the tool collect information about the occasion of the purchase? | To identify whether or not an app collects observable data relating to the occasion on which food was purchased. Occasion can be operationalized as to the circumstances under which food purchase took place. This may include, meals times (i.e., breakfast) or a celebratory occasion (i.e., birthday). | Yes<br><br>No<br><br>No Information |

Regarding the criteria about consumer unit in table 6; three main levels were stated. The levels were based on Geissler & Powers (2011) and FAO/WHO (1996) regarding dietary consumption: 1) Estimations of population food consumption (e.g., domestic food production plus imports and minus exports). 2) Estimations of household food consumption (e.g., food purchases, larder stocks, gifts minus wastage). 3) Estimations of individual food consumption (prospective or retrospective). Data on the national and household level may provide important information such as differences in food availability among different


This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 654280.

communities, or tracking dietary changes in the total population (FAO/WHO, 1996). Since the criteria are based on dietary consumption, it is not as useful as criteria to the data generated by food purchase tools.

Regarding what was purchased, and also how much, when, where, and how, it was recorded at either an act or just an intention, or if it could be both? The theory of planned behaviour (TPB) maintains that behaviour is directly influenced by one's decision to act (i.e., intention) and the control one perceives one has over the behaviour; intention to act, in turn, is dependent on attitudes toward the act, subjective normative pressure to act, and perceived behavioural control (Bagozzi & Dholakia, 2006). For the phases of pre-purchase including levels of information search and increased knowledge & understanding and also planning & organizing purchases (not only making a purchase) data generated from consumers searching for prices or specific products can be of interest even though there was not an actual purchase at that particular time, there were an intention to purchase.

Food purchase data can be recorded in various units, in more ways than by the actual content in energy or nutrient composition of the food. For the purchase services it varies in level of detail about the food/product that is bought and in what variables or "units" it is being recorded. For financial tools the unit is exclusively ´money´ meaning the financial expenditure. The terms ´food´, ´food group´ and product name´ was used when the purchase was detailed as described in table 7 below;

**Table 7. Explanation of variable descriptions for "what is the consumer unit the purchases have been measured?"**

| **´food´** | **´food group´** | **´product name´** | **´ingredients`** |
|---|---|---|---|
| e.g. Apple | e.g. Fruit and vegetables | e.g. Golden Delicious | |
| Cheese | Dairy products | Philadelphia | Sometimes possible to see when there is a ´product´ registered. |
| Bread | Bakery | Kings mill *"great everyday"* | |

As mentioned above, it is important to know whether the tools extend over the phases of the planning of a purchase, consumers increasing knowledge about food, restaurants and stores (before and after a purchase) as well as generating data about the actual point-of-sale. Therefore, criteria had to include about not only what method used to record data, but also what method and units were used for recording how much, when, where and how it was purchased. The most common method used to generate data for how much is the volume or weight of a food product. However, for purchases made in cafés, restaurants and/or take away food units such as serving and portions sizes are more common.


This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 654280.

**Table 8. Purchase variables and descriptions – inputs for "what, where, how and when"**

| **Criteria** *(sub-criteria)* | **Variables** | **Variable Description** |
|---|---|---|
| *Methods: What was purchased?* | *Voice input* | *Speech recognition which do speech-to-text processing, for example adding food items to a shopping list by voice recording* |
| | *Order confirmation* | *A written confirmation of an order which is sent by the producing or service rendering company that accepts the order placed by a purchasing customer.* http://en.ecommercewiki.info/glossary/o/order_confirmation |
| | *Loyalty card* | *A loyalty program for customers of retail businesses, with which shoppers can collect points for purchases and redeem them for vouchers, goods or money.* https://en.wikipedia.org/wiki/Loyalty_program |
| | *Manual input* | *Manual input by the consumer, for example creating grocery shopping list by manually typing/selecting, manual uploading of photos or manual reporting of spending* |
| | *Manual search* | *The user can search manually, for example by categories, product, typing in product/brand names etc.* |
| | *Scanned receipt* | *Scan, capture and store feature for receipts for example to track bank balances* |
| | *Financial transaction* | *Event which involves money or payment, such as the act of depositing money into a bank account or buying goods.* http://www.businessdictionary.com/definition/financial-transaction.html |
| | *Purchase history* | A record of purchases which a customer has made in the past. http://www.dictionarycentral.com/definition/purchase-history.html |
| | *Barcode scanning* | *A barcode reader (or barcode scanner) is an electronic device that can read and output printed barcodes. Ability to scan a barcode to add to a shopping list or look up product information* |
| | *Spectroscopic analysis* | *Analysis of a light spectrum (using a spectrometer) to determine characteristics of its source; for example, analysis of the optical spectrum of a product to determine its composition* |
| | *No information* | *Not enough information to determine the what method* |
| *What unit: purchase* | *Food box* | *A food box, which is delivered to the consumers home address, contains the pre-measured ingredients needed, along with recipe cards* |
| | *Food* | *The type of food which has been purchased such as bread, cake, or steak* |
| | *Food group* | *A collection of foods that share similar nutritional properties or biological classifications* https://en.wikipedia.org/wiki/Food_group *such as dairy products or bakery.* |
| | *Product* | *The specific food product which has been purchased such as Dunkin Donuts, Coca Cola or a Quaker Oats cereal* |
| | *Ingredients* | *Edible substances that are used in a dish or a product* |
| | *Dish* | *A dish is a specific food preparation, a distinct article or variety of food, with cooking finished, and ready to eat, or be served* |
| | *Beverage* | *Any one of various liquids for drinking, usually excluding water* |
| | *Money* | *Any item or verifiable record that is generally accepted as payment for goods* |
| | *No Information* | *Not enough information to determine the what unit method* |


This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 654280.

| | | |
|---|---|---|
| *What is the consumer unit: purchase?* | *Group* | *The aggregated food purchase of an entire group or population* |
| | *Household* | *The collective purchase of food by people occupying the same house or a separate housing unit.* |
| | *Individual* | *The purchase of a single individual* |
| | No Information | *Not enough information to determine the what consumer unit* |
| *Methods: How much was purchased?* | *Serving size* | *A standardized amount of a food such as a cup or an ounce, used in providing information about a food within a food group, such as in dietary guidance. U.S. Department of Agriculture & U.S. Department of Health and Human Services (2010) Dietary Guidelines for Americans, 2010. Washington, DC: U.S. Government Printing Office.* |
| | *Portion size* | *The amount of a food served or consumed in one eating occasion. A portion is not a standardized amount, and the amount considered to be a portion is subjective and varies. U.S. Department of Agriculture & U.S. Department of Health and Human Services (2010) Dietary Guidelines for Americans, 2010. Washington, DC: U.S. Government Printing Office.* |
| | *Product volume* | *The amount of a product that is enclosed within a container e.g. litre, centilitre or millilitre.* http://www.oxforddictionaries.com/definition/english/volume |
| | *Product weight* | *The amount or quantity a product weighs.* http://www.dictionary.com/browse/weight |
| | *Spending* | *For primary finance apps: Money that you spent or plan to spend on activities you enjoy, entertainment, personal things, food etc.* http://dictionary.cambridge.org/dictionary/english/spending-money |
| | No Information | *Not enough information to determine the how much method* |
| *Time unit: Purchase* | *Date* | *The date when the food has been purchased* |
| | *Weeks* | *The food has been purchased within certain weeks* |
| | *Exact time* | *The exact time when the food has been purchased* |
| | *Day Periods* | *The food has been purchased on certain day periods such as morning, or evening* |
| | *Days* | *The food has been purchased on certain days* |
| | *Months* | *The food has been purchased within certain months* |
| | *Years* | *Food has been purchased within a certain year* |
| | *No Information* | *Not enough information to determine the what time unit* |
| *Location unit: purchase* | *No Information* | *Not enough information to determine the where unit* |
| | *Venue Name* | *The name of the venue or restaurant the food has been purchased* |
| | *Geo - Coordinates* | *Units of a coordinate system that enables every location on the earth to be specified by a set of numbers or letters, or symbols* |


This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 654280.

| | | |
|---|---|---|
| *Method: How was it purchased?* | *In-store* | *Happening or existing inside a shop, or available for customers to use or buy inside a shop* http://dictionary.cambridge.org/dictionary/english/in-store |
| | *Online shopping* | *Electronic shopping; shopping done via the Internet; also called e-shopping.* http://www.dictionary.com/browse/online-shopping |

### 3.1.3 TECHNICAL GOVERNANCE QUALITY CRITERIA

Several criteria which are relevant for assessing the quality related to the technical governance of the consumer generated food purchase data were identified. These criteria reflect the widely accepted and recommended FAIR data principle (see e.g., Wilkinson et al., 2016). However, because of the limitations with the exercise, focus remained on those FAIR data principles that do not require an examination of the data structure of the tool or data access documentations in detail.

Accessibility of data refers to how easy it is to access data and metadata (e.g. Dufty, Bérard, Lefranc & Signore, 2014) including the technical infrastructure (e.g. API) for data access (e.g. Dedeke, 2000) and whether data is retrievable using an open, free, and universally implementable communications protocol (e.g. REST) and is represented in a formal, accessible, shared, and broadly applicable language (e.g. JSON; e.g. Wilkinson et al., 2016). In addition to a standardized data access, the protocol should also allow for an authentication and authorization procedure (e.g. OAuth 2.0; e.g. Wilkinson et al., 2016).

The operationalization of the Technical Quality Criteria in RIMS for 'data accessibility' can be seen in Table 9. Firstly, it is important to ascertain whether the data is accessible. The criteria ´data accessible' seeks to answer the question 'is the data collected by the tool accessible directly via the tools infrastructure (not via integrated aggregators)?´ Further accessibility criteria aim to identify whether the tool has any accompanying access documentation, and whether there is a URL to this documentation. The criteria also aims to identify whether the tool has documentation concerning the terms under which the data can be accessed and whether there is a URL to this documentation that users can access. Furthermore, it is an important indicator of data quality that the data can actually be accessed and the form that this access to take (e.g. Email export, web feed, web API), also whether this data can be accessed using a commonly used access protocol.

In order to increase user-friendliness, tools provide technical assistance and support regarding the use of the tool and access to the date, these services require availability of contact information, contact address, and concise and comprehensive documentation of the tool and data access protocols including terms and conditions (e.g. Kim, Eng, Deering & Maxfield, 1999).

Software portability refers to the number of operating systems and devices the application supports (e.g. IOS, Android, Windows; e.g. Meulendijk, Meulendijks, Jansen, & Numans, 2014).


This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 654280.

**Table 9. Technical governance quality criteria**

| **Criteria** *(sub-criteria)* | **Scientific Question** | **Criteria Description** | **Variables** |
|---|---|---|---|
| Is data accessible | Is the data collected by the tool accessible directly via the tools infrastructure (not via integrated aggregators)? | To identify whether or not the consumer generated data collected by the tool is accessible either directly via the tool itself, or via its associated infrastructure (e.g., an API). | Yes<br>No<br>No Information |
| *Access documentation* | *Does the tool provide access documentation?* | *To identify whether or not the owner of the data has provided written documentation instructing users on how to access the consumer generated data associated with the tool.* | *Yes*<br>*No*<br>*No Information* |
| *Terms of access* | *Does the tool provide a term of access document?* | *To identify whether or not the owner of the data has provided written terms by which a user may or may not be permitted to access the data.* | *Yes*<br>*No*<br>*No Information* |
| *URL Terms of access* | *Add URL to terms of use of the data access* | *Provide the URL, if available, to direct users to the terms of access associated with the tool.* | *Text Entry* |
| Implements access protocols? | Can the data be accessed using a commonly used access protocol? | | Yes<br>No<br>No Information |
| *Types of access protocols* | *What commonly used protocol must be implemented to access the data?* | | *Text Entry* |
| *Data formats* | *In what format is the data accessible?* | *To identify the format in which the user generated consumer data is accessible to the user (e.g., Excel, PDF, CSV).* | *Text Entry* |
| Authentication | Does access require authentication? | To identify whether or not access to the data set require the user to be authenticated. That is, some form of validation process is required to authenticate the identification of the user. This may be in the form of a user account held with the data owner. | Yes<br>No<br>No Information |
| Price | Does data access require payment? | To identify, whether or not the owners of the data require a fee/subscription or some other form of payment to access the data. | Yes<br>No<br>No Information |


This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 654280.

### 3.1.4 LEGAL GOVERNANCE QUALITY CRITERIA

Based on the literature research on existing frameworks for the evaluation of eHealth and mHealth tools, several criteria were identified as relevant for assessing the quality related to the legal governance of the consumer generated food purchase data, see table 10. There is a requirement that all tools cover data ownership and data privacy in their licensing agreement, which the consumer accepts at the time of initial use (e.g. Cummings, Borycki & Roehrer, 2013; Adhikari, Richards & Scott, 2014; Blenner et al., 2016). Perceived lack of consent due to data acquisition and usage may undermine public trust (e.g. UNCE, 2014).

Data privacy refers to the disclosure of all data a tool (or any in-app advertiser) collects or accesses on user devices and the applied methods and technology (automatically or manually; e.g. Boulos, Brewer, Karimkhani, Buller & Dellavalle, 2014). This includes collection, storage, and network transmission of user generated data including personal identifiable data and whether the data is securely encrypted during and after those workflows (e.g. Nije, 2013a), and the duration and termination of data storage (e.g. Cummings, Borycki & Roehrer, 2013).

In addition, data privacy refers to the (secondary) usage of the user generated data such as making data accessible to the general public or sharing data with other affiliated or unaffiliated third-parties such as analytics and advertising services, or data brokers (e.g. Nije, 2013b; Cummings, Borycki & Roehrer, 2013).

Data ownership refers to both the possession of and responsibility for information. Ownership implies power as well as control. The control of information includes not just the ability to access, create, modify, package, derive benefit from, sell or remove data, but also the right to assign these access privileges to others (Loshin, 2002). Loshin (2002) identifies a list of parties laying a potential claim to data such as the party that creates or generates the data (e.g. the app user), the enterprise in which the data is created (e.g. the app vendor) or the individual or organization that buys or licenses data (e.g. third parties and business partners).

Data privacy and ownership may have a significant influence on the intended use of the data given legal limitations, organizational restrictions, confidentiality and privacy concerns (e.g. UNCE, 2014).

Data security refers to the extent to which access to information is restricted appropriately to maintain its security (e.g. by authentication; e.g. Knight & Cowan, 2005; Schulze & Kromker, 2010; Martinez-Perez, de la Torre-Diez, Candelas-Plasencia & Lopez-Coronado, 2013). Data security can be assessed on several levels such as data-level, application-level, network-level, and host-level security (e.g. Ho, Lee & Armstrong, 2013). In addition, data security can refer to the data storage such as local storage versus cloud-based storage or the availability of data backup systems (e.g. Ho, Lee & Armstrong, 2013).


This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 654280.

**Table 10. Legal governance quality criteria**

| **Criteria** *(sub-criteria)* | **Scientific Question** | **Criteria Description** | **Variables** |
|---|---|---|---|
| Terms of use | Does the tool provide terms of use document? | To establish whether or not the tool provides a statement or legal documentation that discloses the terms under which the consumer may use the tool. | Yes<br>No |
| URL Terms of use | Add URL to the terms of use document | To provide a URL to a website (or similar) on which the documentation relating to the terms under which the consumer may use the tool are found. | Text Entry |
| *Privacy Policy* | *Does the tool have a privacy policy document?* | *To establish whether or not the tool provides a statement or legal document that discloses some or all of the ways in which the company responsible for the tool gathers, uses, discloses and/or manages the tool users data.* | *Yes*<br>*No* |
| URL Privacy Policy | Add URL to the privacy statements document | To provide a URL to website (or similar) on which a statement or legal documentation that discloses some or all of the ways in which the company responsible for the tool gather, uses, discloses and/or manages the tool users data can be found. | Text Entry |
| *Data ownership* | *Who holds the ownership of the user generated data (user content)?* | *To identify the individual and/or company that has legal rights and control over a single piece or set of data generated through consumer use of the tool.* | *Text Entry* |
| Data usage vendor | Does the tool vendor retain the right to access and exploit the user generated data (publish, distribute, publically display)? | To establish whether or not the vendor of the tool (e.g., app store), retains the right to access and exploit the consumer generated data produced by the tool. | Yes<br>No |
| *Personal information* | *Does the tool collect personal identifiable information (e.g., during registration)?* | *To identify what, if any, personal identifiable information about the user of the tool is collected during the registration process.* | *Text Entry* |
| Informed consent | Is the user asked permission about collecting personal identifiable information? | To establish whether or not the tool service user has granted permission for the data they generate to be used by another party in the full knowledge of the possible consequences. | Yes<br>No<br>No information |


This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 654280.

| | | | |
|---|---|---|---|
| Types of Personal information | What types of personal identifiable information does the tool collect? | To identify the types of personal identifiable information collected by the tool about the tool user. | Text Entry |
| Public profile | Does the tool create a public profile of the users' personal data? | To establish whether or not the tool creates a public profile of either the consumers' personal identifiable information, or consumer generated data collected via the tool. | Yes<br>No<br>No Information |
| Device Data | Does the app collect device data after installation/visit? | To establish whether or not the tool is collecting data as to the device via which the consumer uses the tool. | Yes<br>No<br>No Information |
| Type of device data | What type of device data does the tool collect? | To identify the type of information that is being collected about the device via which the consumer uses the tool (e.g., IP address). | Text Entry |
| Cookies | Does the homepage/website of the tool store cookies on a user's computer? | To establish whether or not the homepage/website of the tool stores cookies (data sent from the website to the user's browser). | Yes<br>No<br>No Information |
| Web Beacons | Does the homepage/website of the tool store web beacons to track the online moments of users? | To establish whether or not the homepage/website of the tool uses web beacons (embedded objects that invisibly check whether the user has accessed content). | Yes<br>No<br>No Information |
| Identifiable data sharing (Affiliated parties) | Will collected personal identifiable data be shared with affiliated third parties (with confidentiality agreements)? | To establish whether or not personal identifiable information collected about the consumer with be shared with an affiliated third party, either with or without the consent of the consumer. | Yes<br>With consent<br>No<br>No Information |
| Identifiable data sharing (Unaffiliated parties) | Will collected personal identifiable data be shared with unaffiliated third parties (without confidentiality agreements)? | To establish whether or not the personal identifiable information collected about the consumer will be shared with unaffiliated third parties, either with or without the consent of the consumer. | Yes<br>With consent<br>No<br>No Information |
| Data storage | Where does the system store the data it generates? | To identify the storage system on which the owner of the data stores the consumer generated data collected by the tool. | Web server storage<br>Device storage<br>No information |


This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 654280.

| | | | |
|---|---|---|---|
| Data deletion | Is the user able to delete or ask for deletion of his or her personal identifiable information (e.g., after account termination)? | To establish whether or not the consumer is able to delete, or request deletion, of his or her personal identifiable information collected via the tool. | Yes<br>No<br>No Information |
| Usage Analytics | Does the homepage/website of the tool use third-parties for advertising and usage analytics? | To establish whether or not the homepage/website of the tool uses third-party advertising on their homepage/website, and/or a third-party for the analysis of homepage/website usage. | Yes<br>No<br>No Information |
| Third party services | Does the tool provide any third party services? | To establish whether or not the tool uses any third-party services. That is, are any services provided by the tool outsources to a company, or individual, other than the tool owner. | Yes<br>No<br>No Information |
| Social Network Sharing | Can the collected data be shared with social networks? | To establish whether or not the tool has the facility for the consumer to share their data collected within the tool with a social network (either their own social network, or that of another). | Yes<br>No<br>No Information |
| Data Encryption | Does the tool encrypt the collected data? | To establish whether or not the tool encrypts the consumer generated data. That is, are the data converted into another form which cannot easily be understood by anyone other than an authorized party? | Yes<br>No<br>No Information |

# 4. Summary and discussion

As mentioned, the overall aim of RICHFIELDS is to design a Research Infrastructure for the collection, integration, processing and sharing of consumer generated data as related to food behaviour and lifestyle determinants, because of that it is crucial to give structure around what kind of consumer generated data that is out there.

The work behind this deliverable was not only to list a series of quality criteria, but to visualize the potential opportunities with the consumer generated food data and how that reflects on food intake activities, seen from in this particular case - food purchase data, also to identify factors influencing the quality of the data and also highlight gaps and needs with it as part of the collection, integration and dissemination process. The quality criteria, as set out in this deliverable, aim to assess whether the consumer generated data has the potential for use in research (through individual tools such as smartphone apps, websites, and sensors).


This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 654280.

WP7 has been the initiator for the creation of the quality criteria currently found in RIMS. These criteria are based on aspects of health and lifestyle specific to food consumption. Purchase behaviour is different in some aspects and therefore some specific criteria are not as relevant for food purchase data. Technical and legal governance are of highest relevance for the research questions within RICHFIELDS, however these are areas which are difficult to interpret for non-experts. When it comes to the current selection of technical criteria, it is still very much open to discussion regarding whether these are appropriate criteria, as, the field for tools like these are changing rapidly with a potential risk that we are not focusing on the most optimal criteria.

As for the recordable data and the result of the inventory to be presented in deliverable 5.1, it will also be hard to do the inventory of the tools to that level of detail as the quality criteria suggests. For many tools, it is not possible to respond to these criteria, particularly with the feasibility parameters worked to in this exercise. That is to say, it is not possible to easily identify certain aspects of a tool's quality without either expert knowledge of the fields of ICT and Law, and without the downloading and testing of a tool, the examination of a tool's data structure and/or the examination of a hosting data infrastructure. This is therefore a potentially time consuming and costly process to validate the quality of consumer generated data produced via a tool.


This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 654280.

# 5. References


Abbar, S., Mejova, Y., Weber, I. (2014) You Tweet What You Eat: Studying Food Consumption Through Twitter, CoRR abs/1412.4361

Adhikari, R., Richards, D., & Scott, K. (2014). Security and Privacy Issues Related to the Use of Mobile Health Apps. Paper presented at the 25th Australasian Conference on Information Systems.

Bagozzi, R.P. & Dholakia, U.M. (2006) Antecedents and purchase consequences of customer participation in small group brand communities. *Intern. J. of Research in Marketing 23 (2006) 45–61.*

Blenner, S. R., Kollmer, M., Rouse, A. J., Daneshvar, N., Williams, C., & Andrews, L. B. (2016). Privacy Policies of Android Diabetes Apps and Sharing of Health Information. *Jama, 315*(10), 1051-1052.

Boulos, M. N., Brewer, A. C., Karimkhani, C., Buller, D. B., & Dellavalle, R. P. (2014). Mobile medical and health apps: state of the art, concerns, regulatory control and certification. *Online J Public Health Inform, 5*(3), 229.

Cummings, E., Borycki, E. M., & Roehrer, E. (2013). Issues and considerations for healthcare consumers using mobile applications. *Stud Health Technol Inform, 183*, 227-231.

Dedeke, A. (2000). *A Conceptual Framework for Developing Quality Measures for Information Systems.* Paper presented at the Conference on Information Quality Retrieved from http://mitiq.mit.edu/ICIQ/Documents/IQ%20Conference%202000/Papers/AConceptualFramework4DevelopQualityMeasure.pdf

Dufty, D., Bérard, H., Lefranc, S., & Signore, M. (2014). *A Suggested Framework for the Quality of Big Data*. Retrieved from http://www1.unece.org/stat/platform/download/attachments/108102944/Big%20Data%20Quality%20Framework%20-%20final-%20Jan08-2015.pdf?version=1&modificationDate=1420725063663&api=v2

FAO/WHO. (1996). *Preparation and use of food-based dietary guidelines*.

Fried, D., Surdeanu, M., Kobourov, S., Hingle, M., Bell, D. (2014). Analyzing the language of food on social media, in IEEE International Conference on Big Data. https://arizona.pure.elsevier.com/en/publications/analyzing-the-language-of-food-on-social-media

Geissler & Powers (2011)

Howard, JA. & Sheth, JN., (1969). A Theory of Buyer Behavior, New York, Wiley. p. 25f

Ho, B., Lee, M., & Armstrong, A. W. (2013). Evaluation criteria for mobile teledermatology applications and comparison of major mobile teledermatology applications. *Telemed J E Health, 19*(9), 678-682.

Kim, P., Eng, T. R., Deering, M. J., & Maxfield, A. (1999). Published criteria for evaluating health related web sites: review. *BMJ, 318*(7184), 647-649.



This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 654280.

Knight, S., & Cowan, E. (2005). Developing a Framework for Assessing Information Quality on the World Wide Web. *Informing Science Journal, 8*.

Kroski, E. (2008). What Can You Do with the Mobile Web? Mobile Web Applications. In *On the Move with the Mobile Web: Libraries and Mobile Technologies.*

Köster, E.P. (2009). Diversity in the determinants of food choice: A psychological perspective. *Food Quality and Preference 20 (2009) 70–82*

Levy, M., Weitz, B., & Grewal, D. (2014). Retailing Management, 9th Edition. New York, McGraw Hill education

Loshin, D. (2002). Knowledge Integrity: Data Ownership. *Data Warehouse magazine (Online) June 8, 2004*. Retrieved from www.datawarehouse.com/article/?articleid=3052

Lupton, D. (2016). The diverse domains of quantified selves: self-tracking modes and *dataveillance. Economy and Society Volume 45 No. ,1 pp. 101–122*

Martinez-Perez, B., de la Torre-Diez, I., Candelas-Plasencia, S., & Lopez-Coronado, M. (2013). Development and evaluation of tools for measuring the quality of experience (QoE) in mHealth applications. *J Med Syst, 37*(5), 9976.

Mejova, Y., Haddadi, H., Noulas, A., Weber I. (2015) #foodporn: Obesity patterns in culinary interactions, in Proceeding DH '15 Proceedings of the 5th International Conference on Digital Health, pp 51-58.

Meulendijk, M. C., Meulendijks, E. A., Jansen, P. A. F., & Numans, M. E. M. (2014). *What concerns users of medical apps? Exploring non-functional requirements of medical mobile applications.* Paper presented at the Twenty Second European Conference on Information Systems, Tel Aviv Retrieved from http://aisel.aisnet.org/cgi/viewcontent.cgi?article=1004&context=ecis2014

Njie, C. M. L. (2013a). *Technical Analysis of the Data Practices and Privacy Risks of 43 Popular Mobile Health and Fitness Applications*. Retrieved from http://www.privacyrights.org/mobile-medical-apps-privacy-technologist-research-report.pdf

Njie, C. M. L. (2013b). *Privacy Aware Mobile Application Development*.

Okumus, B. & Bilgihan, A. (2014). Proposing a model to test smartphone users' intention to use smart applications when ordering food in restaurants. *Journal of Hospitality and Tourism Technology* 5 No. 1, 2014 pp. 31-49.

Saarijärvi, H., Mitronen, L. & Yrjölä, M. (2014) From selling to supporting – Leveraging mobile services in the context of food retailing. *Journal of Retailing and Consumer Services* 21: 26-36.

Schroeck, M., Shockley, R., Smart, J., Romero-Morales, D., & Tufano, P. (2012). *Analytics: The real-world use of big data*. London: IBM Global Business Services Business Analytics and Optimisation in collaboration with Säid Business School, University of Oxford. Retrieved from http://www-935.ibm.com/services/us/gbs/thoughtleadership/ibv-bigdata-at-work.html

Schulze, K., & Kromker, H. (2010). *A Framework to Measure User Experience of Interactive Online Products.* Paper presented at the 7th International Conference on Methods and Techniques in


This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 654280.

Behavioral Research. Retrieved from http://measuringbehavior.org/files/ProceedingsPDF(website)/Schulze_FullPaper5.3.pdf

Shankar, V., Venkatesh, A., Hofacker, C. & Naik, P. (2010) Mobile Marketing in the Retailing Environment: Current Insights and Future Research Avenues. *Journal of Interactive Marketing* 24 (2010) 111–120

Sharma S., De Choudhury, M. (2015) Detecting and Characterizing Nutritional Information of Food and Ingestion Content in Instagram, in Proceeding WWW '15 Companion Proceedings of the 24th International Conference on World Wide Web, pp 115-116

Sobal, J. (1991) Obesity and socioeconomic status: A framework for examining relationships between physical and social variables. *Medical Anthropology*, **13**, 231–247.

Sobal, J., Bisogni, C. A. & Jastran, M. (2014), Food Choice Is Multifaceted, Contextual, Dynamic, Multilevel, Integrated, and Diverse. Mind, Brain, and Education, 8: 6–12.

Solomon, M. R., Russel-Bennett R. & Previte, J. 2013. Consumer behaviour: Buying, having, being. 3rd edition. Frenchs Forest: Pearson Australia.

UNCE, 2014

Velasco, E., Agheneza, T., Denecke, K., Kirchner, G., Eckmanns, T. (2014).Social media and internet-based data in global systems for public health surveillance: a systematic review Milbank Quarterly 92, 7

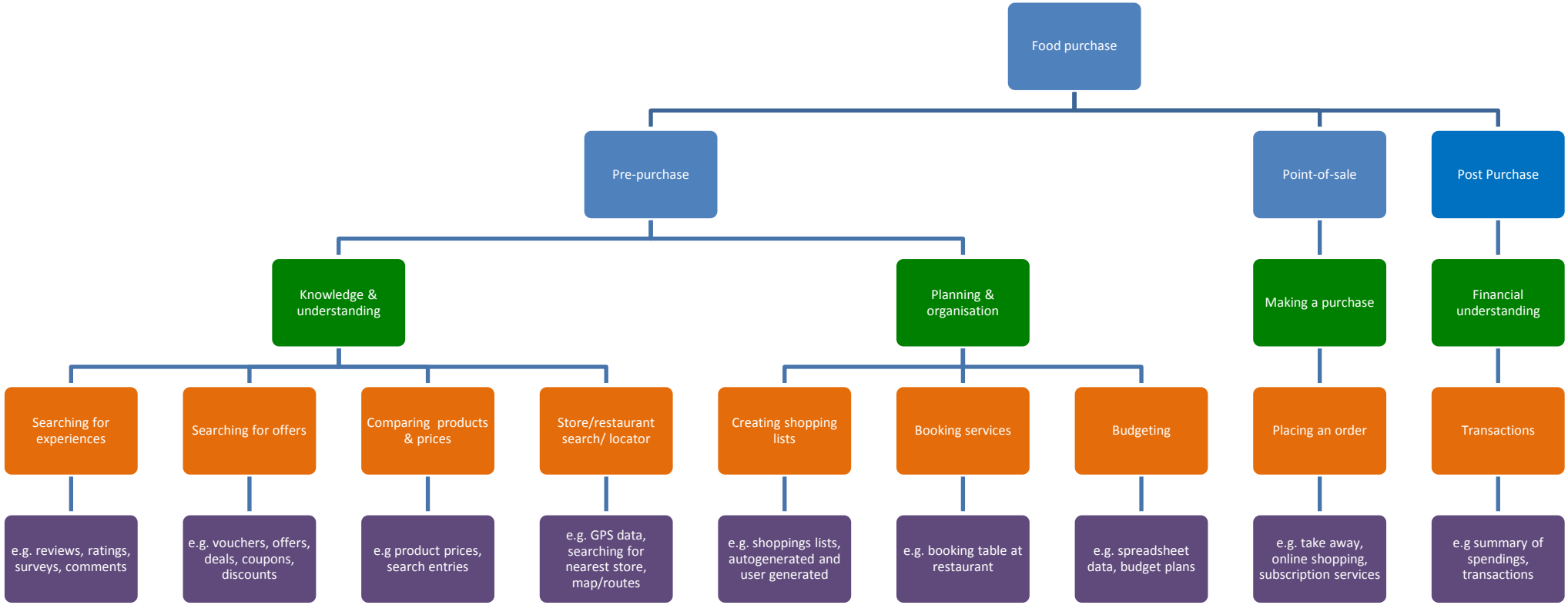
Weber, I., Achananuparp, P. (2016) Insights from machine-learned diet success prediction, in PSB, 2016. http://arxiv.org/abs/1510.04802

Wilkinson, M. D., Dumontier, M., Aalbersberg, I. J., Appleton, G., Axton, M., Baak, A., . . . Mons, B. (2016). The FAIR Guiding Principles for scientific data management and stewardship. *Sci Data, 3*, 160018.



This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 654280.

# Appendix 1 – typology





This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 654280.