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

## RICHFIELDS Working Package 3
Deliverable D3.3

## Report from first Stakeholder workshop

**Date delivered: M20**

**Authors:**
Kerry A. Brown, Lada Timotijevic (USurrey - UK)
Siân Astley (EuroFIR AISBL – BE)
Paul Finglas (IFR – UK)

**Deliverable lead beneficiaries:**
University of Surrey (UK)
EuroFIR AISBL (BE)


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:** | Report from first Stakeholder workshop |
| **Deliverable No.:** | D3.3 |
| **Authors:** | Kerry A. Brown, Lada Timotijevic, Siân Astley (EuroFIR), Paul Finglas (IFR) |
| **Reviewer:** | Karin Zimmermann – Project Coordinator<br>Pieter van 't Veer – Scientific Coordinator |
| **Start date:** | 1.10.2015 |
| **Delivery date:** | 30.05.2017 |
| **Due date of deliverable:** | 30.11.2016 |
| **Dissemination level:** | PU |
| **Status:** | Final |

**Change history:**

| Version | Notes | Date |
|---|---|---|
| 001 | D3.3 Report from first Stakeholder workshop DRAFT | 30.10.2016 |
| 002 | D3.3 Report from first Stakeholder workshop UPDATED | 30.03.2017 |
| 003 | D3.3 Report from first Stakeholder workshop SUBMITTED | 30.05.2017 |

**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.

## Table of Contents

Table of Contents ..... 3
1. Executive Summary ..... 5
2. Background ..... 7
2.1 RICHFIELDS objectives ..... 7
2.2 RICHFIELDS structure ..... 8
2.3 RICHFIELDS Phases 1-2: Findings to date ..... 8
B.3.1 Phase 1 ..... 8
B.3.2 Phase 2 ..... 9
2.4 User requirements analysis ..... 9
2.5 Information architecture ..... 10
2.6 RICHFIELDS final design ..... 10
3. Workshop objectives and outputs ..... 11
3.1 Aims of the workshop ..... 11
3.2 Output of the workshop ..... 12
4. Workshop methodology ..... 13
4.1 Recruitment and participants ..... 13
4.2 Procedure and materials ..... 14
5. Results from the workshop ..... 15
5.1 Measurement ..... 15
5.2 Inference ..... 16
5.3 Problems and potential solutions ..... 17
5.3.1 Ownership of the activities at a consortium level ..... 17
5.3.2 Clarity of vision and offerings ..... 18
6. Workshop conclusions ..... 20
6.1 WP3 Conclusions on the process and outcomes ..... 20
6.2 WP3 Conclusions on stakeholder reflections ..... 20
6.3 Consequences for RICHFIELDS going forwards ..... 20
6.4 Schedule for Workshops 2 and 3 and the second Platform ..... 20


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

Annex 1: Workshop attendance ..... 21
A1.1 Invitations ..... 21
A1.2 External participants ..... 22
A1.3 RICHFIELDS beneficiaries ..... 26
Annex 2: Workshop programme ..... 31
A2.1 Workshop agenda ..... 31
A2.2 Workshop presentations ..... 32
A2.2.1 10:30 – 10:35 Welcome (Siân Astley, EuroFIR AISBL - BE) ..... 32
A2.2.2 10:35-10:45 Introduction (Paul Finglas, IFR - UK) ..... 33
A2.2.3 10:45-11:05 Introduction: Phase 1 (Monique Raats, Uni Surrey – UK) ..... 35
A2.2.4 10:45-11:05 Introduction: Phase 2 (Bent Egberg Mikkelsen, Aalborg Uni – DK) ..... 38
A2.3 Workshop materials ..... 47
A2.3.1 Summary of the B2R - GfK Denmark ..... 47
A2.3.2 Summary of the G2R - Statistics Denmark ..... 51
A2.3.3 Paprika recipe manager ..... 52
A2.3.4 Pantelligent ..... 54
A2.3.5 OCADO ..... 56
A2.3.6 MyfitnessPal ..... 58
A2.3.7 Lloyds bank mobile banking ..... 61
A2.3.8 Fitbit ..... 64
A2.4 Sign-in sheets ..... 67
Annex 3: Workshop notes: presentation and breakout session ..... 71
A3.1 Group assignment and questions ..... 71
A3.2 Measurement break-out session feedback ..... 72
A3.3 Inference break-out session feedback ..... 77
3.4 Plenary discussion ..... 83
Annex 4: Workshop evaluation ..... 89


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

## 1. Executive Summary

Stressing the need for world-class research infrastructures, EU Horizon 2020 has provided financial support for RICHFIELDS, which commenced on 1st October 2015 for three years. LEI Wageningen UR (NL) coordinates the project, bringing together 16 organisations from 12 countries. Competences include nutrition, sociology, information management, ICT, business and consumer science.

RICHFIELDS aims to design a consumer-data platform that will collect and connect information about food behaviours, specifically the determinants of food intake. To achieve this, analysis of user requirements is essential to ensure the consumer data platform is fit-for-purpose. Activities to characterise RICHFIELDS end-users and stakeholders and their requirements are being conducted iteratively alongside design of the platform, and includes informal interviews, questionnaires, inventories, and workshop discussions with user groups/stakeholders in Phases 1-2.

WP3 will deliver three Stakeholders' workshops during the lifetime of the project, the first of which was held on 27th September 2016 at Schiphol Amsterdam (NL). The aim of this workshop was to support on-going work specifying and characterising datasets providing information about consumer behaviour around food choices. A secondary goal was to secure a group of individuals with expertise not represented amongst the beneficiaries (e.g. App developers) or potential users (e.g. researchers) who could engage with RICHFIELDS to support design of the platform. Results from this workshop will feed directly into RICHFIELDS activities, informing the architecture and user requirements analyses.

The external delegates had expertise in food composition, nutritional surveys and statistics (Aida Turrini – IT), sports nutrition (Paolo Colombani – CH), App development, sales and marketing (James Lay – UK), data quality (Pieter Francois – UK), applied use of consumer data (Roel van der Heijden – NL), data analysis (Giulia Vilone – IE), research infrastructure ethics, data protection and intellectual property (Maud Alligier – FR, ECRIN), computer systems programming (Paul Allington – UK) and social simulation (Kristrún Gunnarsdóttir – UK). In addition, 12 beneficiaries were represented. Following the welcome and introductory presentations, delegates participated in two group sessions and plenary feedback. Participants were pre-assigned to groups with a good mix of disciplines including both beneficiaries and external stakeholders. Examples of data sources supported discussions around the extent to which data can be used to understand better consumer behaviours around purchase, preparation and consumption.

Two issues were identified as important for RICHFIELDS with respect to data, namely measurement and inference, which were the focus of this workshop. Definition of appropriate vocabularies/ ontology for a common understanding of metadata and data type and quality, however, were not achieved, largely because of the lack of clarity around the final design at this early stage in the project. Nevertheless, exchange with the stakeholders was very useful to shape potential core offerings (tools and services), and stakeholders are interested in the concept of RICHFIELDS. There was agreement about advantages and disadvantages of the various data sources, meta-data needed to utilise potential resources, potential difficulties in understanding/ interpretation of data, and exploitation of existing and future data sources within a research framework


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

Several issues for WP3 and appropriate solutions to improve external uptake and internal engagement have been identified, and will be applied for workshops 2 and 3 in 2017. Specifically, beneficiaries need to be supported to engage actively with WP3 events through the creation of small teams responsible for development of the programme and delivery of the event, and these individuals should be drawn from the Project Management Team as well as the WPs involved, directly or indirectly.

Following this workshop, RICHFIELDS must decide the exact nature of the proposed consumer data platform, specifically whether it is a platform exclusively for the use of researchers or if there will be personalised feedback to consumers, potential data sources (i.e. individuals or tools and services, such as Apps), likely cooperation of providers and whether and how this can be secured, and the types and quality of data and meta-data resources to be included. This will help clarify aspects of the design for stakeholders. These decisions will come from on-going WP activities and outcomes will form the basis of a core offering at the minimum viable product level. Similarly, deciding on the nature of the platform will enable outputs (e.g. vocabularies/ ontology) to be defined using tools such as focus groups and pilot studies, if not by RICHFIELDS then in the future as activities in the framework of the ESFRI roadmap.


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

## 2. Background

RICHFIELDS aims to design a consumer-data platform that will collect and connect information about food behaviours, specifically the determinants of food intake. The first two phases (Phase 1 WP5-7; Phase 2: WP8-10) will deliver knowledge about consumer-related data (e.g. type, quality, restrictions etc.), which might be linked to the RICHFIELDS platform. Outputs from these phases will then be used to support design of the platform during the last phase of RICHFIELDS (Phase 3: WP11-13).

To-date, the first two phases of RICHFIELDS have identified sources of un-/structured data that could provide information about the determinants of consumer dietary intake. These sources are diverse, including data collected and provided through consumer mobile phone applications, commercial company sales, national/ government resources, research infrastructures and research facilities.

Throughout Phases 1-2, each data source is being considered for the:

1. **Scientific case** - why the data source should be linked to create the RICHFIELDS platform (e.g. what research questions can be asked/ answered with these data or what can be inferred/deduced from these data?)
2. **Ability to comply with FAIR[1] principles** for data sharing including whether data are findable, accessible, interoperable or re-usable (e.g. what are the technical and legal constraints of accessing, processing, linking or disseminating certain data sets, such as the constraints from intellectual property or data protection standards?)
3. **User strategy** (e.g. what are the characteristics of the RICHFIELDS platform end-users and what are the user requirements of the data/ tools/ services that could be provided by RICHFIELDS?)

### *2.1 RICHFIELDS objectives*

RICHFIELDS will design the technical requirements for a consumer data platform to collect and connect, compare and share information about our food behaviours. We seek to determine which facilities, resources, and services could support research activities to learn more about what we choose to eat, and how and why we make those choices. RICHFIELDS is exploring the integration of data on food purchase, food preparation and food consumption, generated from different sources:

- Consumers (e.g. apps, sensors)
- Business (including retail, e-commerce e.g. sales)
- Research (including European and International research, e.g. surveillance data)

A business model will outline the services provided by the RICHFIELDS platform, and how these will generate revenue to sustain it in the longer-term, while a roadmap will outline the steps needed to introduce a platform that can serve the whole of Europe.

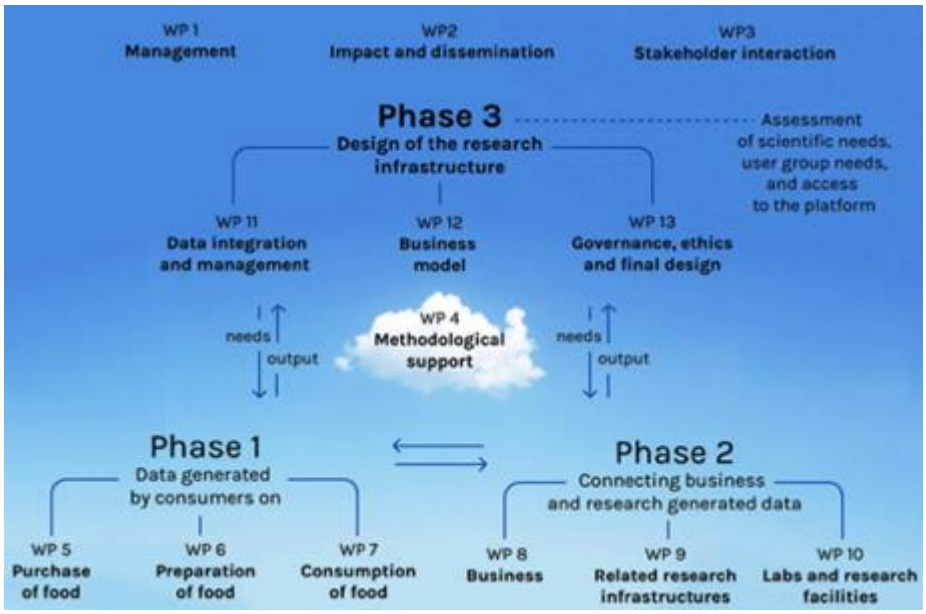
[1] FAIR principles https://www.force11.org/group/fairgroup/fairprinciples


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

### *2.2 RICHFIELDS structure*

Stressing the need for world-class research infrastructures, EU Horizon 2020 has provided financial support for RICHFIELDS, which commenced on 1st October 2015 for three years. LEI Wageningen UR (NL) coordinates the project. Sixteen organisations from 12 countries, bring together competences including nutrition, sociology, information management, ICT, business, consumer science, and food processing. The first two phases of RICHFIELDS (Phase 1 WP5-7; Phase 2: WP8-10) will deliver in-depth knowledge about the available consumer-related data and, based on these outputs, the future requirements for such a platform (design) will be developed (Phase 3: WP11-13) (Figure 1).

**Figure 1. RICHFIELDS structure**




### *2.3 RICHFIELDS Phases 1-2: Findings to date*

#### B.3.1 Phase 1[2]

An inventory management system (RIMS) has been created for storage and assessment of an online inventory of tools (e.g., mobile phone applications), which produce consumer generated food and/ or beverage purchase, preparation or consumption data. RIMS is comprised of two parts: (1) a typology categorising the purpose of tools and (2) metadata to enable assessment of data quality, either related to a scientific case (e.g. are the data sufficient to answer a what/ who/ why/ how/ where research question) or whether the data are findable, accessible, inter-operable or re-useable (e.g. legal, governance or technical data management constraints). Information about these is fundamental to developing the architecture and governance structure of the RICHFIELDS platform.

[2] D5.3, 6.3, 7.3: Susanne Ekman, Anouk Geelen Naomi Klepacz, Marcus Maringer, Anne Normann, Anne Normann, Muriel Verain


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

### B.3.2 Phase 2[3]

Case studies in work packages 8-10 allow a more detailed approach to investigate the technical components, interfaces and services necessary for data to be linked to create a functioning RICHFIELDS platform. These case studies include:

- Work package 8: Three case studies addressing business generated data on purchase and procurement: (i) Coop DK, (ii) Statistics DK, (iii) Göteborgs Stad SE
- Work package 9: Four case studies exploring the potential for delivering data and content to the RICHFIELDS platform from existing infrastructures or those currently under development: (i) food composition and food attributes (EuroFIR, FoodExplorer, ePlantlibra, Brandbank, FoodWiz); (ii) Standardised food intake from population based surveys (Globodiet); (iii) Clinical interventions; and (iv) consumer diet, health and lifestyle (PRECIOUS, Quisper).
- Work package 10: Three case studies investigating laboratories and facilities that undertake consumer research on food choice, purchase and consumption: (i) the Fake Food Buffet at ETH Zurich (food choice); (ii) the FoodScape Lab at Aalborg University (food choice, consumption); (iii) Restaurant of the Future at Wageningen University (food choice, purchase and consumption).

Phases 1-2 both reflect on the scientific case for using the data source, compliance with FAIR principles and user strategy (as described in Section 2: Background).

## *2.4 User requirements analysis*

An on-going task throughout RICHFIELDS will be a user requirements analysis. A series of tasks will be performed to characterise RICHFIELDS end-users and stakeholder groups their requirements to ensure the platform is fit-for-purpose. User requirements analysis will be conducted iteratively alongside the design of the RICHFIELDS platform. The user requirements analysis tasks include:

- Informal interviews with stakeholders at the RICHFIELDS stakeholder platform (2nd June 2016, Brussels – BE, see D3.1 for meeting details)
- Questionnaire survey distributed to existing research infrastructures on user groups and provision of food and health research services to establish provisional user group profiles
- Phase 1-2 research activities including inventories, focus groups and workshop discussions with user groups/stakeholders in Phases 1-2. This information will inform the research questions in the survey below.
- Questionnaire survey and follow-up interviews with user groups/ stakeholders to identify user groups and user requirements
- Workshops and stakeholder platform meetings throughout Phases 1-3.

[3] Phase two protocol: Paul Finglas, Sophie Hieke, Haris Hondo, Bent Mikklesen, Kwabena Ofei, Mark Roe


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

## 2.5 Information architecture[4]

Work package 4 will collate information generated during Phases 1-2 and the user requirements work performed to produce a RICHFIELDS information architecture draft. This will outline the key principles and building blocks for RICHFIELDS to aid the development of the final design in Phase 3.

## *2.6 RICHFIELDS final design*[5]

Phase 3 will use the knowledge generated in Phases 1-2 as well as any additional Phase 3 activities to generate three elements of the final RICHFIELDS platform design:

1. Semantic model - this is necessary to encode data and information, and allow the sharing (re-use) of data with various RICHFIELDS end-users or information systems (software agents). Work package 11 aims to produce an ontology and set of classes to aid the re-use and integration of data, information and knowledge.
2. Business model - work package 12 will produce different business models dependent on the value proposition (service offered), supply chain configuration (means to deliver services to users) and revenue system (remuneration mechanism for the platform).
3. Governance model – will be depend on how governance is defined, i.e. which elements of governance will be included within the design of RICHFIELDS. Issues related to FAIR data, such as data ownership, privacy, intellectual property rights, and ethics will all need to be considered.

[4] Led by WP4
[5] WP11-13 internal notes 2015: Marc-Jeroen Bogaardt; Elisabetta Chierici, Barbara Koroušić Seljak


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

## 3. Workshop objectives and outputs

### *3.1 Aims of the workshop*

The aim of this workshop was to support the on-going work regarding requirements for specifying and characterising the wide range of datasets identified as providing information about consumer behaviour around food choices. The workshop aimed to secure a group of individuals with expertise not represented amongst the beneficiaries (e.g. App developers) or potential users (e.g. researchers) who could engage with RICHFIELDS to support design of the consumer data platform.

Results from this workshop should feed directly into RICHFIELDS activities, informing the architecture and user requirements analyses. To achieve this, two issues were identified as important for RICHFIELDS to link with data, and create tools and services related to dietary behaviour (e.g. purchase, preparation, food choice, consumption), namely:

1. **Measurement** (units of observation), which is associated directly with the scientific case and compliance with FAIR principles as well as user strategy
2. **Inference** (interpretation): What questions might analysis answer regarding dietary behaviour (e.g., purchase, procurement, preparation, food choice, consumption) (user strategy)?

Measurement, characterising the observations in a way that enables inferences about the phenomenon being studied, requires that the following issues be addressed:

**1. What is the core unit of observation linked to the data?**

The unit of observation is the major entity being analysed (who, what, when, where, why), and clarity about these elements is essential to understand the limits for inferences. Understanding what is measured (e.g. sampling specification) provides the basis for interpretation.

In relation to the RICHFIELDS scope, the following units of observation are possible:

- Individual (captured through the measures of behaviour, cognitions, emotions);
- Household (captured through the measures of e.g. energy consumption, purchase data, sensors);
- Organisation (captured through the purchase/procurement data, legal/documentary data),
- Location (which may include a range of spatial data from the GPS-recorded movement data, through to epidemiological data on e.g. morbidity and mortality within a specific geographic location, waste production etc.);
- Social networks (articulated social networks – e.g. Facebook links and behavioural social networks, derived from communication patterns and cell coordinates)

**2. What are the attributes ascribed to the observation to make it measurable?**

For RICHFIELDS, how we can define the attributes and use them to infer from the data might be easier in certain domains (e.g. frequency of intake) than in others (e.g. recipe sharing). Equally, some offer more significant (e.g. self-reported individual consumption) in understanding determinants of food behaviours than others (e.g. purchase at household level).


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

Inference is interpretation based on the analysis of data, and clarity is needed about what can be inferred from data captured or rather the extent to which these data reflect research concepts, such as attitudes, cognitions, emotions, etc. in food behaviours. Extracting meaning from data enables factors (e.g. purchasing behaviour) to be transformation into variables that form the basis of analysis. Questions to be addressed include:

**a. To what extent can the data be interpreted in a way that enables extrapolation of meaning based on current core constructs of consumer behaviour? What data can we analyse and what questions can such analysis can answer? What is the unit of analysis?**
Answers link back to scientific research questions and forward to decisions rules for inference. For instance, in accessing an individual's daily food intake data it is important to understand whether aggregate intake data permit claims about population level behaviour.

**b. What criteria need to be in place to enable inferences?**
Sampling is essential for social science, as it defines the unit of analysis, which allows inferences about validity (i.e. context), ethics, reliability, etc. but these issues also enable understanding of uncertainty, biases and limitations of the data and our understanding.

### *3.2 Output of the workshop*

Outputs from the workshop were anticipated to be, amongst other things, definitions of appropriate vocabularies/ ontology for a common understanding of:

1. Metadata
2. Data type and quality inventory (Phase 1) as related to un-/structured sources (Phase 2)

Outputs from the workshop will be summarised as criteria that should be considered as part of the design (Phase 3), and provide a case for the scientific need and potential for links across datasets.


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

## 4. Workshop methodology

### *4.1 Recruitment and participants*

The aims of the first Stakeholder Platform were to (1) introduce RICHFIELDS to the stakeholder community and (2) identify the stakeholder needs on the scientific and business cases. This was achieved with a meeting on 2nd June 2016 (Brussels – BE), which is described in D3.1 Report from first Stakeholder Platform meeting. A parallel aim of the larger platform was to identify individuals with expertise not represented amongst the beneficiaries (e.g. App developers) or potential users (e.g. research) who could engage with RICHFIELDS to support design of the research infrastructure in a series of workshops, the first of which is described herein (Amsterdam Schiphol – NL, 27th September 2016) and focussed on measurement of consumer data and inference.

Thus, potential participants were selected from those invited to the first Stakeholder Platform who has expressed an interest in working more closely with RICHFIELDS in a workshop environment, individuals who were from business and industry, technology, research or funding bodies identified in a stakeholder analysis (see Annex 1: Workshop attendance, 1.1 Invitations for more information) including, for example, existing research infrastructures, and personal connections across the consortium as well as beneficiaries, either as speakers or representing specific sectors.

From these sources, 40 individuals with expertise in 10 areas (economics, epidemiology, public health, app developers, big data curators and analysts, data users for research, diet and health researchers) were identified as having a specific, measurable, achievable, relevant and timely interest in the design of RICHFIELDS. Invitations were sent to these individuals during June-September 2016, and nine accepted (28%, see A1.2 External participants).

Briefly, external participants included individuals with expertise in food composition, nutritional surveys and statistics (Aida Turrini – IT), sports nutrition (Paolo Colombani – CH), App development, sales and marketing (James Lay – UK), data quality (Pieter Francois – UK), applied use of consumer data (Roel van der Heijden – NL), data analysis (Giulia Vilone – IE), research infrastructure ethics, data protection and intellectual property (Maud Alligier – FR, ECRIN), computer systems programming (Paul Allington – UK) and social simulation (Kristrún Gunnarsdóttir – UK).

In total, 22 beneficiaries were invited and 32 external delegates. Of these, 12 (55%) representatives of beneficiaries attended (see A1.3 RICHFIELDS beneficiaries). Briefly, Moniqe Raats, Lada Timotijevic, Charo Hodgkins, Kerry A. Brown – (University of Surrey – UK, Phase 1-3 and expertise in consumer behaviour), Paul Finglas (IFR, UK -diet and health research), Siân Astley (EuroFIR, BE - diet and health research, and communications), Angelika Mantur (EuroFIR BE - dietetics and nutrition), Naomi Klepacz (University of Surrey, UK – psychology), Marcus Maringer (WUR, NL - psychology) and Barbara Koroušić Seljak (JSI, SI - computer science and informatics). The University of Surrey was over-represented because of staff changes (Kerry A. Brown – finished, 30th September 2016, Charo Hodgkins – started, 1st September 2016).


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

### *4.2 Procedure and materials*

Following the welcome and introductory presentations (see A2.1 Workshop agenda and A2.2 Workshop presentations), there were two group sessions (90 minutes each) followed by 30 minutes' feedback in plenary session, totalling four hours closing working.

The participants were assigned to two groups of seven with a good mix of disciplines, internal and external stakeholders. The groups were given five 'case studies' (see A2.3 Workshop materials) that collect data that might be useful for determining purchase, preparation and consumption. One case also contained information about business-to-reseller (B2R) and government-to-reseller (G2R) data.

It total, there were eight 'case studies', tools used by consumers, namely GfK Denmark (B2R) (purchase of organic food and beverages), Statistics Denmark (G2R) (grocery purchase at the household level), Paprika recipe manager (recipes), Pantelligent (cooking), Ocado (grocery shopping), MyfitnessPal (health and wellbeing), Lloyds Bank mobile banking (personal finances), and Fitbit (health and wellbeing). There are tens of thousands of such Apps, and related tools and services, collecting data from consumers that could be relevant in understanding purchase, preparation and consumption behaviours. These were selected as good examples, based on number of users, familiarity and functionality (i.e. bringing together different aspects of purchase, preparation and consumption or other relevant behaviours).

In the first breakout session, the groups discussed understanding the nature of consumer-generated data, and reflected specifically on what data from the examples would allow potential users to do within their disciplines, what research questions could the data help answer, what unit(s) of observation were represented by these data and who or what could these data might be attributed to (e.g. individual, household, an organisation, geographical location, social interaction, etc.), the shortcomings of these data, and how useful these data would be to a (discipline) studying food-related phenomena. In the second breakout session, the groups discussed unpinning meta-data, specifically what it is necessary to know about the to make it useful for research of food related phenomena (e.g. intake, practices, health) and what meta-data are essential for understanding.

Within each group, there was one facilitator and one rapporteur, and the groups were asked to record as much of the discussion as possible, using a template provided in Word and/ or hard copy (see 3.2 Measurement break-out session feedback).

The groups spent 10 minutes summarising their findings prior to feedback.


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

## 5. Results from the workshop

The aim to secure a group of individuals with expertise not represented amongst the beneficiaries (e.g. App developers) or potential users (e.g. research) who could engage with RICHFIELDS to support design of the research infrastructure in a series of workshops was achieved. In future, we aim to retain these individuals and continue to work with them and others in workshops 2 (April 2017) and 3 (Late 2017), before the final Platform in (March-April 2018) and the final conference (September 2018), where the design will be presented for the first time.

### 5.1 Measurement

In summary, based on the case studies (8), the groups collectively identified the following with respect to data:

⇒ ***What do these data allow you to do within your discipline? What research questions can they help answer (or allows you to ask) within your discipline?***

- Cross-reference with other source allows criteria to be determined (e.g. anthropometric)
- Better knowledge of approaches, e.g. search strategies, recording of data (manual or automatic)
- How much individuals move and comparisons between times
- Language choice and its relation to cultural norms
- Longer-term data will reveal more usual behaviours
- Market share of products
- Media or public health interventions messages are part of the exchange of information
- Messages are not quality controlled (e.g. consumption of sweets containing vitamin C are given the same positive reinforcement as fruit containing vitamin C)
- Purchasing behaviour might reflect or impact supply chain and demand
- Seeking information (e.g. exercise regimens) cannot be assumed to translate into actions
- Wearables and apps are related to exercise <u>and</u> diet, not just exercise <u>or</u> diet

⇒ ***What unit of observation does it represent – who or what can you attribute to these data to – the individual, household, an organisation, geographical location, social interaction etc.?***

- Household level (but it might not be the only retailer used)
- IP Address
- User community

⇒ ***What are the shortcomings of the data?***

- App developers' willingness to share, terms and conditions for users to share with third parties
- Bias, uncertainty, falsification (deliberate or neglect)
- Ownership
- Children not included
- Data inputs are all different, data types not known
- Ethical and privacy issues; opt in or opt out; consent versus informed consent
- Household and population-wide purchase behaviours not individuals
- Individual activity or consumption not household and population
- Demographics of users; hard-to-reach, low socioeconomic groups not users
- List of ingredients can be structured but not processing (e.g. foods might be grilled or fried)


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

- Lack context of activities (purchase, preparation or consumption)
- Public and private sharing behaviours
- Relationship between purchase and consumption is unclear
- RICHFEILDS does not want to collect and store these data only access on-demand
- Short-termism (during dieting and or fitness behaviours)

⇒ ***How useful would these data be for studying food-related phenomena?***

- Aggregate level, can access data individually or collectively
- Food purchase, depending on the type of family, food preparation changes
- Link to other sources to provide aggregated data

The wider discussions and comments with respect to data and specific case studies are included in A3.2 Measurement break-out session feedback (Group 1 and Group 2).

### 5.2 Inference

In summary, based on the case studies (8), the groups collectively identified the following with respect to inference:

⇒ ***What do we need to know about data to make these useful for research of food related phenomena (e.g. intake, practices, health)?***

- Community/ population
- Context of input
- Cooking practices
- Data quality
- Impact of updates on data and users
- Legislation, regulation, ethical issues
- Measures
- Research question/ framework
- Numbers of users (consumers) and or sources (Apps)
- Processing of raw data, if any
- Product identity
- Standardisation in reporting of data and activities
- Uncertainty
- Validity of the data (objective measures)

⇒ **What do we need to be recording about the data in a meta-data document?**

- Source
- Organisation of raw data
- Context (e.g. how much of the behaviour is captured)
- Applicability to individuals and populations
- Aggregated or raw data; processing if aggregated


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

The wider discussions and comments with respect to inference and specific case studies are included in A3.3 Inference break-out session feedback (Group 1 and Group 2).

An issue identified during plenary discussion was accessing of data directly from individuals, which would allow motivation to be considered (e.g. healthier living) and include potentially relevant meta-data (e.g. anthropometric measures, geographic location) against indirect access (from a provider). The value of individual (direct) data in terms of population outcomes, which is more important for public health, has yet to be established. Also, there is a widespread move away from one-size-fits-all to more personalised information/ advice, which raises two issues that RICHFIELDS would have to manage: (1) expectation that, in exchange for data, individual receive personalised advice and (2) RICHFIELDS could impinge on existing tools, commercial and research, that already engage with individuals with this intent.

Similarly, accessing consumer information indirectly would increase the volume of data available, reducing the impact of intra-individual differences allowing functional inter-individual differences to be revealed and exploited. However, this approach raises ethical and commercial issues. Where individuals might expect personalised information/ advice, App developers would have an expectation of commercial gain from sharing data over and above what they might achieve independently, bearing in mind the primary goal from the majority is income generation/ profit not improving public health. Also, terms and conditions might have to be adapted to allow them to share some or all their customers' data, subject to opt-in or opt-out choices. A third option would be to collect data without the knowledge of users; this is already done where individuals leave wi-fi and or Bluetooth option broadcasting, and activities (shopping in a closed space, such an airport) are tracked based on IP address. This might have role in, for example, supermarkets but there are ethical and technical issues in extending this approach to food behaviours beyond simply purchase.

These discussions also highlighted different needs of potential providers and users in terms of data. For example, a sports nutritionist, diet and health researcher and a commercial enterprise were all interested in the impact of intake on health, as measured using biomarkers associated with specific health-related endpoints/ risk. In comparison, programmers and curators sought clean, aggregated data with clearly defined ontologies, regardless of their original purpose. Meta-data are primarily needed to validate data quality and support adaptive-ness. RICHFIELDS would need to know how sources collect meta-data as well as what descriptors are being recorded. These issues would also impact any potential feedback to users in terms of recommended changes in behaviour. It was suggested that, rather than discuss what might be needed, RICHFIELDS should undertake a pilot study with selected Apps, creating and API to automatically or semi-automatically define the meta-data. This will be considered by WP11 Data integration & data management, as part Task 11.2: Semantic data model of the RI Consumer Data Platform.

See Annex 4: Workshop evaluation for feedback regarding the content and delivery of the workshop.

### *5.3 Problems and potential solutions*

#### *5.3.1 Ownership of the activities at a consortium level*

One concern during the planning of this workshop was the extent of buy-in, at the consortium level, and engagement with the development of the content, objectives and outputs. This does not reflect an


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

unwillingness to interact with stakeholders, as these activities are occurring in some WPs independent of WP3, but rather the perception that these workshops are a WP3 activity. In fact, the WP3 workshops were included in the Description of Action to provide opportunities for the consortium as-a-whole, Phases 1-3 and individual WPs, together or independently, to interact with stakeholders without the burden of organising meetings.

This perceived issue will be addressed through the lessons learned, specifically:

- The role of WP3, to ensure the RICHFIELDS platform design is optimised for a range of users through building and maintaining effective interaction with stakeholders throughout the life of the project and beyond, will be emphasised at project meetings and during planning of subsequent activities
- The objectives of WP3, to establish a vibrant and active stakeholder Platform to engage with the project and work proactively with stakeholders through a series of related workshops, and how these activities are intended to help guide beneficiaries in the RICHFIELDS platform design will be promoted proactively amongst internal stakeholders (e.g. WP-leaders, phase-leaders, project management board) since this cannot occur without wider participation (ownership)
    - WPs 5-13 benefit directly from the activities of WP3 and this will be addressed with WP-leaders in the lead up to the second stakeholder workshop (4th April 2017), which will also be led by WP4/ University of Surrey (UK) and facilitated by WP3/ EuroFIR AISBL (BE)
    - The project management board will be engaged directly to support development of the content to assure added value in the second stakeholder workshop alongside those occurring in individual WPs
    - Activities and proposed content for discussion of second stakeholder workshop will be presented at the consortium meeting in March 2017 to encourage involvement and increased perceived value

### *5.3.2 Clarity of vision and offerings*

An issue that arose during discussions with stakeholders, during both the larger platform and this workshop, was the lack of apparent clarity in the vision for RICHFIELDS, specifically what would be offered in terms of tools and services. Potential data providers (e.g. programmers) and users (e.g. researchers) struggled to understand how they might use this consumer data platform over other tools or services, particularly in the absence of a specific research question or commercial drivers. More specifically, programmers and developers wanted to understand more about what was wanted to determine how data should be collected and ensure outputs are relevant whilst social and biomedical researchers expressed the need for a research framework to define what data are valuable and what meta-data are needed to give it context. However, whilst biomedical researchers favoured a specific research question before interrogating the data, social researchers were more interested in the wider landscape and what behaviours the data might inform with pre-assumptions.

In bringing the pre-design phase to stakeholders, there is always a risk that the outcomes are not clear, and the specific needs and goals of potential data providers and users would not be addressed immediately at the individual or even level of different specialisms. Whilst this is an important consideration for workshops 2 and 3, and the final Platform, it is also one that will resolve itself, as the


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

consortium works through the various phases and Deliverables generate outputs that can be translated into tools and services, based on robust ICT, business models and ethics and governance. By engaging stakeholders at this early stage, however, even if the outcomes are not entirely clear, the feedback is useful, helping to identify potential issues and shape the discussions across the WPs. Overall, therefore, the outcomes from workshop 1 – real and implied – have been useful in shaping RICHFIELDS's core offering at the minimum viable product level.


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

## 6. Workshop conclusions

*6.1 WP3 Conclusions on the process and outcomes (WP leader, Siân Astley – EuroFIR AISBL, BE)*

- Attracting stakeholders to events hosted by RICHFIELDS to discuss the design is achievable
- Beneficiaries need to be encouraged and supported to engage actively with WP3 events
- There is a perceived lack of clarity in the vision for RICHFIELDS that could weaken stakeholder participation, but should resolve as the project progresses
- Outputs from workshop 1 were not fully achieved, specifically definition of appropriate vocabularies/ ontology for a common understanding of (a) metadata and (b) data type and quality
- Exchange with the stakeholders is useful to shape potential core offerings (tools and services)

*6.2 WP3 Conclusions on stakeholder reflections*

- Stakeholders with a variety of expertise are interested in the concept of RICHFIELDS
- Broadly, there is agreement amongst stakeholders about (1) advantages and disadvantages of potential data sources, (2) meta-data needed to utilise these resources, (3) potential difficulties in understanding and exploitation existing and future data sources within a research framework.

*6.3 Consequences for RICHFIELDS going forwards*

- RICHFIELDS must decide the exact nature of the proposed consumer data platform with some degree of urgency, specifically whether (1) it is a platform exclusively for the use of researchers or if there will be personalised feedback to consumers, (2) data sources will be individuals or tools and services, such as Apps, (3) cooperation of App providers can be secured or if Apps to collect data need to be developed and (4) types and quality of data and meta-data resources to be included. Currently, the lack of clarity around this is confusing and alienating all stakeholders, internal and external, and risks jeopardising their continued participation in development of the design. These decisions will come from on-going WP activities and outcomes will form the basis of a core offering at the minimum viable product level.
- Defining the nature of the platform, as above, will enable outputs (e.g. vocabularies/ ontology) to be defined using tools such as focus groups and pilot studies, if not by RICHFIELDS then in the future as activities in the framework of the ESFRI roadmap.
- Internal stakeholders need to have a sense of ownership over WP3 events. WP3 and WP4 will work with WP2, the management team and WP-leaders to encourage participation.

*6.4 Schedule for Workshops 2 and 3 and the second Platform*

- **Workshop 2 Tuesday 4th April 2017, MAI 40 rue Washington, 1050 Brussels, BE**
    - December 2016 – draft agenda
    - January 2017 – identify and invite delegates
    - February 2017 – publish agenda and draft supporting information
    - March 2017 – Open registration, secure numbers for catering, etc.
- **Workshop 3 October-November-December 2017 – to be agreed**
- **Platform 2 mid-late April 2018 – to be agreed**


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

## Annex 1: Workshop attendance

*A1.1 Invitations*

Between November 2015 and June 2016, individuals and organisations including umbrella organisations (i.e. SME Associations and Research Infrastructures, e.g. ECRIN) were identified as RICHFIELDS stakeholders from a range of sources. This process created two lists: (1) RICHFIELDS STAKEHOLDERS Contact Details containing 269 individuals or individuals representing an organisation and (2) RESEARCH INFRASTRUCTURES containing 187 individuals representing an EU-funded research infrastructure, based on information provided in http://bit.ly/2eU9gFZ, with a likely interest in or potential relationship with RICHFIELDS (i.e. biomedical and or social science).

Review and update of these lists is on-going and will be throughout the lifetime of the project. However, information includes (stakeholder) name, contact person, contact form URL or email address(es), website, country, city and category of interest group, specifically whether research, technology, funding or industry. Additional information for each research infrastructures is acronym, project/ overarching organisation and purpose.


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

*A1.2 External participants*

**Aida Turrini - *CREA-Alimenti e Nutrizione (IT)***

Aida's expertise is mainly in nutritional surveys, acquired over 30 years of work at the research institute, now Research Center for Food and Nutrition (CREA - Council of Agricultural Research and Economics) formerly the National Institute of Nutrition. She has a statistical background and has developed research nutritional database systems as well as food coding and classifications. She is author of peer-reviewed papers, co-authored as a research group member, and reviewed proceedings, chapters and books. Overall, Aida has 98 international publications (60 peer-reviewed as author) and 59 national publications (38 peer reviewed as author). She is registered as ORCID 0000-0002-2188-9406 and RESEaRCH ID K-5353-2016. She is also an experience teacher in graduate and post-graduate courses. Aida is collaborating in national (e.g., CLUSTER AGRIFOOD) and international committees, such as the Network on Food Consumption Data (European Food Safety Authority), and international association like EuroFIR AISBL. Currently, she is a senior researcher (level I) leading the research group in "Nutritional food consumption population study". Her main research task is the coordination of the fourth Italian nationwide dietary survey.

**Paolo Colombani - *Independent consultant (CH)***

Paolo Colombani is a nutrition scientist. He studied food engineering at the ETH Zurich (MSc) and did his PhD on nitrogen metabolism in endurance athletes at the same university (1993-1998). For 15 years Paolo lectured and carried out research projects in the areas of physical activity, nutrition and health and was head of the Swiss food composition database for six years. He was partner of the FP6 Network of Excellence EuroFIR and president of EuroFIR AISBL. In 2010, Paolo started delivering scientific support in nutrition as an independent consultant to the food industry, top management of different industries including banks, elite athletes, Swiss Olympic, Antidoping Switzerland and many more. Today, he is self-employed and continues to deliver scientific support in food and nutrition. He founded the Swiss Sports Nutrition Society and he is his current president. As a partner of an US based technology start-up, he is also strongly involved in the field of personalised/precise recommendations across different health areas.


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

**James Lay - *Food Angels UK Ltd. (UK)***

I am Managing Director of Food Angels UK Ltd. a partners of the Institute of Food Research on a European project regarding the eating habits of 16 and 17 year old adolescents. We wrote the software and provided the database for an app. similar to MyFitnessPal for the project. I have a background in sales and marketing fence I am a Fellow of the Institute of Sales and Marketing.

**Pieter Francois - *University of Hertfordshire and University of Oxford (UK)***

Pieter Francois is a Senior Lecturer in Digital History, University of Hertfordshire, and the research coordinator of the Cultural Evolution Lab, at the Institute of Cognitive and Evolutionary Anthropology, University of Oxford. In 2011, he co-founded the Seshat: Global History Databank project, which aims to be the premier home to test social sciences theories with historical and archaeological data. The Seshat project makes full use of RDF technology and of the possibilities the Semantic Web offers. The development of the Seshat platform is funded primarily through the Horizon2020 project 'ALIGNED quality-centric, software and data engineering' of which Dr. Francois is a Co-PI. His key research interest lies in how data quality is defined and measured in large collaborative projects.

**Roel van der Heijden - *University Medical Center Groningen (NL)***

After obtaining my MSc. in Medical Biology at the Radboud University Nijmegen, I continued my metabolic studies at the University of Groningen Medical Center (UMCG) where I obtained my PhD studying the role of diet-induced systemic inflammation in obesity and linked micro- and macro vascular pathologies. Having left the lab, currently, I work at the UMCG's Center for Development and Innovation as innovation officer Food & Health. In this role, I'm brokering between UMCG researchers and industrial parties in launching innovative projects in different EU funding schemes (EIT-Health, H2020). At the national level UMCG is coordinator of the Dutch node for DISH-RI, aiming to establish a food and health research infrastructure in the Netherlands under the DISH-RI EU umbrella coordinated by the WUR. For RICHFIELDS especially, the expertise present in Groningen at the level of consumer science (citizens and patient) and large research and data infrastructures could be of major interest.


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

**Giulia Vilone - Creme Global Ltd. (IE)**

Giulia Vilone is a senior data analyst at Creme Global Ltd., a data science company based in Dublin (IE) specialising in predictive modelling and software for assessing consumer health.

**Maud Alligier - European Clinical Research Infrastructure Network (ECRIN)**

Maud completed her PhD in 2011 (University of Lyon, FR) in the field of Nutrition, Metabolism and Endocrinology. She is in charge of the scientific coordination of the FORCE network (French Obesity Research Network of Excellence). This network includes more than 20 centres (Clinical and research units and Investigation platforms) and aims to develop and conduct multicentre clinical trials. To achieve this objective, Maud works to ensure the harmonisation of clinical investigation practices and data collection/storage/analysis between the various centres. Alongside this on-going project, Maud is closely involved with ENPADASI, which aims to deliver an open access research infrastructure that will contain data from a wide variety of nutritional studies, ranging from mechanistic/ interventions to epidemiological studies, including a multitude of phenotypic outcomes that will facilitate combined analyses in the future. In ENPADASI, Maud helps the Pr Laville with the coordination of the Regulations (WP5), which deals with Ethics, Data Protection and Intellectual Property of the data sharing. Recently, Maud has also become a part-time (20%) member of the ECRIN nutrition team, as a project manager. ECRIN nutrition hub is a structured European human nutrition research centres network (80 centres in 21 European Countries) dedicated to performing clinical research in Nutrition with the highest standards of quality, following European and International Directives for Clinical Research.


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

**Paul Allington - The Code Guy (UK)**

I am a Microsoft certified professional developer, systems architect and problem solver, with 11 years continuous experience both as an employee working for start up companies and as an entrepreneur starting up a digital agency, growing a team of skilled developers, designers and marketing professionals to bring to market innovative ideas, solutions to life in general and solving business problems for my clients. My skills center around .Net and C#, although my degree in Computing for Artificial Intelligence means I have the foundation to work around many languages and techniques.

**Kristrún Gunnarsdóttir - Centre for Research in Social Simulation (CRESS), University of Surrey (UK)**

Kristrún Gunnarsdóttir is a Research Fellow at the Centre for Research in Social Simulation (CRESS). She is the project manager of HomeSense, a three-year ESRC-funded project to develop tools to facilitate and simplify the use of sensors in social research. The project team includes researchers at the 5G Innovation Centre at Surrey and is in partnership with Microsoft Research Cambridge. Drawing on recent developments in the use of fixed and mobile sensors, the team is trialling the use of a sensor suite in a sample of UK households and will build from that a package of training and capacity-building resources with the UK National Centre for Research Methods (NCRM). Kristrún completed her PhD and MSc at Cardiff University (UK). Her background covers conceptual/ visual ICT designs, IT administration and development, philosophy and sociology. Her research interests draw mainly on Ethnomethodology, activity and cognitive theories, to examine the relationship between computational functions and subtle human judgement in meaning-making, decision-making and ordinary action. She has also worked on a number of EC-funded projects to investigate ethical, legal and socio-economic implications of new-emerging ICTs, applying socio-technical analyses and assessments of ICT/data-driven innovations and visionary work, public engagement, policy and governance programmes.


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

*A1.3 RICHFIELDS beneficiaries*

**Monique M. Raats - Consumer Behaviour and Health Research Centre, University of Surrey (UK)**

Professor Monique Raats is Director of the University of Surrey's Food, Consumer Behaviour and Health Research Centre. Her portfolio of research is wide ranging in terms of topics being addressed (e.g. food choice, food preparation, policy development, food labelling), and methodologies used (e.g. qualitative, quantitative, stakeholder consultation). She has published over 110 peer-reviewed papers, 19 book chapters, and co-edited two books (The Psychology of Food Choice; Food for the Ageing Population). She is a founding member of the International Society of Behavioral Nutrition and Physical Activity. In 2011 Monique joined the UK's Scientific Advisory Committee on Nutrition and is a member of its Subgroup on Maternal and Child Nutrition. Currently she is a partner in the Horizon 2020 RICHFIELDS project that aims to design a consumer-data platform to collect and connect, compare and share information about our food behaviours, to revolutionise research on every-day choices made across Europe and PROSO project that is to providing guidance on how to encourage engagement of citizens and third sector organizations, like non-governmental organizations (NGOs) and civil society organizations (CSOs), in Europe's research and innovation processes. She also coordinates the REDICLAIM project, which investigates how EU legislation impacts on the substantiation and use of "reduction of disease risk" claims on food and drinks.

**Bent Egberg Mikkelsen - Aalborg University (DK)**

**Javier de la Cueva - Independent consultant (ES)**

**Paul Finglas - Institute of Food Research (UK)**

Paul Finglas joined the Institute of Food Research in 1981 and is, currently, Head of the Food Databanks National Capability at IFR (www.ifr.ac.uk/fooddatabanks), and research leader in Food and Health. He has, for most of his science career, been involved in food nutrition and health including food composition and analysis (nutrients & bioactive compounds), traditional and ethnic foods, food description and data quality, dietary intake assessment, nutritional labelling & health claims, reformulation and impact on food intake and health, personalised nutrition and research infrastructures. Paul has considerable experience in both participating in EU projects in food, nutrition and health (PRECIOUS, REFRESH & RICHFIELDS) as well as leading (EuroFIR, TDS-EXPOSURE & BACCHUS). Paul has a broad range of experience in science publishing and is editor for the journals Food Chemistry, and Trends in Food Science and Technology. Paul has a degree in Chemistry from Aston University in Birmingham and has published over 150 publications on a wide range of topics in food science and nutrition. He is also the President for EuroFIR AISBL, a non-profit organisation based in Brussels (BE).


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

**Siân Astley - European Food Information Resource (EuroFIR AISBL, BE)**

Siân has worked extensively with individuals and organisations throughout Europe from a variety of disciplines including research, food and biotech industries and the media. She is author of more than 300 popular science articles for magazines and trade publications as well as 27 peer-reviewed papers, and she was awarded her Diploma in Science Communication in 2009 (Birkbeck University of London). After 14 years as a bench-scientist, Siân became Communications Manager for NuGO, one of the first FP6 Networks of Excellence, and was the European Communications Manager for the Institute of Food Research in Norwich (UK) until April 2012. Currently, she is a senior researcher and the training and communications manager for the European Food Information Resource (EuroFIR AISBL), supporting research as well as training and communications activities within EU-funded research projects and networks. She is also an independent science communicator and an editor for Food Chemistry.

**Angelika Mantur - European Food Information Resource (EuroFIR AISBL, BE)**

Angelika joined EuroFIR in November 2015 as a Research Associate. She is responsible for support with the on-going and new EU projects on food and health, assisting with membership recruitment and training. She studied dietetics at the Medical University of Bialystok (Poland), where she obtained her Master degree in June 2013. She worked as a dietitian for MedFitness, where she was responsible mainly for consultations with clients and nutritional advice. She has participated in many medical and fitness conferences and trade shows, presenting body composition analysers and performing body composition analysis. Angelika took part in 'Keep the Balance', organised by the Polish Dieticians Association and National Food and Nutrition Institute, where for three months she educated patients on nutrition and wellbeing.

**Lada Timotijevic - Food, Consumer Behaviour and Health Research Centre, University of Surrey (UK)**

Having completed my PhD in 2000 (University of Surrey) in the area of identity processes in the context of social and cross-cultural mobility, I have subsequently worked within advertising industry (J. Walter Thompson). I joined the *Food, Consumer Behaviour and Health Research Centre (FCBH)* at the University of Surrey (Department of Psychology) in 2002, a multidisciplinary research centre that brings together skills and expertise from across the University in order to address research questions on food related policy, consumer behaviour and public health. Since my arrival, I have played an instrumental role in the success of the Research Centre, working on research projects of substantive theoretical and applied relevance. I work within the critical public health framework and my empirically-oriented work has focused on understanding the role and nature of public and stakeholder engagement and dialogue in policy and science, risk


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

perception and governance, and science-policy interaction. Policy relevance is a key theme across my research projects, and my work is aimed at both understanding the processes of policy making, and contributing evidence on which to base policies. I am particularly interested in public health nutrition, sustainable diets and illness prevention.

**Kerry Ann Brown - Research Fellow and NHS South East Coast Research Design Adviser, University of Surrey (UK)**

Kerry is a researcher with interests in public health, primarily the development and implementation of dietary guidelines. Kerry's academic training has been in the biological and medical sciences (Sport Science, Nutrition), as well as the social sciences (Psychology). She has over 10 years' research experience working within multi-disciplinary and international teams and designing studies using a variety of methods (qualitative, quantitative and mixed). As a NHS research design adviser, Kerry works with front line clinicians to facilitate research to benefit NHS patients. As a research fellow, Kerry has been involved in regional and national projects related to child and adolescent dietary behaviour; as well as European projects related to micronutrient recommendations (EURRECA), health claims (CLYMBOL), and European research infrastructure (EuroDISH, RICHFIELDS).

**Naomi Klepacz - Consumer Behaviour and Health Research Centre, University of Surrey (UK)**

Naomi Klepacz is a Chartered Psychologist specialising in health and a Research Fellow in the Faculty of Health and Medical Sciences at the University of Surrey. She holds a PhD Health Psychology and her research interests focus on health communication and promotion. She is particularly interested in the role of visual imagery in people's understanding of health and memory for health related information. She has an interest also in health damaging behaviour and psychological interventions used for primary prevention, as well as the processes influencing the effective delivery of health care and its impact on individuals, family members and caregivers. Naomi has experience of conducting research both as part of both UK and EU funded projects, employing both qualitative and quantitative methods.


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

**Charo Hodgkins - Consumer Behaviour and Health Research Centre, University of Surrey (UK)**

Charo Hodgkins is a science graduate and started her career with GSK as a development chemist. In 1997 she moved to the retail sector as Head of Technical Services for Superdrug Stores PLC. During her 14 years in industry, she gained extensive experience of managing technical and data management projects within both branded and retail environments. Her expertise includes research and development, manufacturing, and quality/supply chain management for a wide range of products including, pharmaceuticals, medical devices, foods, toiletries and non-foods. Her responsibilities also involved extensive auditing of production facilities across Europe and the development and delivery of training packages in Continuous Improvement, HACCP, Data management, Crisis Management and Problem Solving techniques. In 1999, Charo took a short career break to start a family and in 2002 joined the Food, Consumer Behaviour and Health Research Centre at the University of Surrey as a Research Fellow. She has been active in a number of UK and EU funded research projects in the area of food, consumer behaviour and public health. Charo has recently completed her PhD investigating the role of food composition data, nutrition information and health claims in communicating healthier food choices.

**Marcus Maringer - Wageningen University (NL)**

Marcus Maringer is a Social Psychologist who acquired his PhD in Behavioral and Social Sciences in 2007 at the University of Groningen. In his research and theoretical approach, he emphasized the socially situated and embodied nature of social cognition and the role of accessible information, feelings and subjective experiences in judgment and decision-making processes. He has a keen interest for innovative scientific methodology especially in the field of computer-aided research. He is an expert in technical solutions engineering for social scientific research on mobile phones and as scientific technical advisor and software engineer he has designed and developed research applications and supporting infrastructures for diverse Institutions including the University Medical Center Groningen, Radboud University or Pluryn Research & Development. In 2015 he joined the Wageningen University and the RICHFIELDS project as a researcher exploring and evaluating publically available technical solutions for collecting food consumption and associated lifestyle data.


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

**Barbara Koroušić Seljak - Institut Jozef Stefan (SI)**

Barbara Koroušić Seljak earned her PhD at the University of Ljubljana in Computer Science and Informatics, and works for the Computer Systems Department, Jožef Stefan Institute, in Ljubljana (Slovenia). Currently, she is the Assistant Professor at the Jožef Stefan International Postgraduate School. She is a member of the Executive Board of the Slovenian Society for Clinical Nutrition and Metabolism as well as of EuroFIR. In the project RICHFIELDS she is the leader of WP11, where a RI platform will be designed considering state-of-the-art ICTs for collecting big and open data created by consumers and researchers or generated by machines, such as sensors gathering information, digital pictures and videos, purchase transaction records, GPS signals, etc.


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

## Annex 2: Workshop programme

### *A2.1 Workshop agenda*

| | |
|---|---|
| **09:30 – 10:30** | **Arrival and registration** |
| **10:30 – 10:35** | **Welcome (Siân Astley, EuroFIR AISBL - BE)**<br>• **Welcome, aims of the meeting**<br>• **Structure of the day** |
| **10:35-10:45** | **Introduction (Paul Finglas, IFR - UK)**<br>• **Overview of RICHFIELDS** |
| **10:45-11:05** | **Introduction**<br>• **Phase 1 (Monique Raats, Uni Surrey - UK, 10 mins)**<br>• **Phase 2 (Bent Egberg Mikkelsen, Aalborg Uni - DK, 10 mins)** |
| **11:05-11:15** | **Coffee break** |
| **11:15 -12:30** | **Group session 1: Understanding the nature of the data**<br>Reflect on the five cases presented and address the following issues:<br>1. What do these data allow you to do within your discipline? What research questions can the data help answer (or allows you to ask)?<br>2. What unit of observation does the data represent – who or what can you attribute these data to – individual, household, an organization, geographical location, social interaction etc.?<br>3. What are the shortcomings of these data?<br>4. How useful would these data be to a (discipline) studying food-related phenomena? |
| **12:30-13:00** | **Plenary session – feedback from the groups** |
| **13:00 – 14:00** | ***Lunch buffet*** |
| **14:00 – 15:30** | **Group session 2: Developing a meta-data**<br>1. What do we need to know about the data to make it useful for research of food related phenomena (e.g. intake, practices, health)?<br>2. What meta-data do we need to be recording? |
| **15:30-15.45** | **Coffee break** |
| **15:45-16:15** | **Plenary Session – feedback from the groups** |
| **16:15-16:30** | **Closing plenary & Good bye**<br>• **Summary of results**<br>• **Next steps and good bye** |

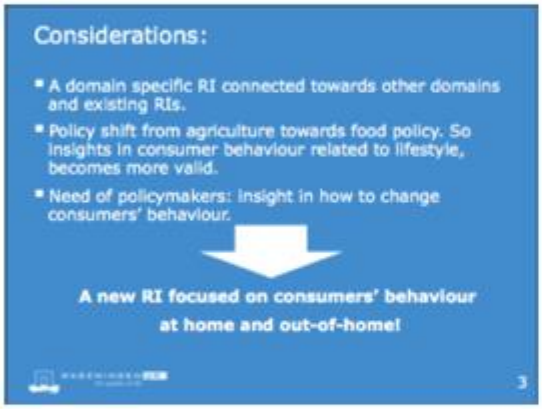

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 654280.

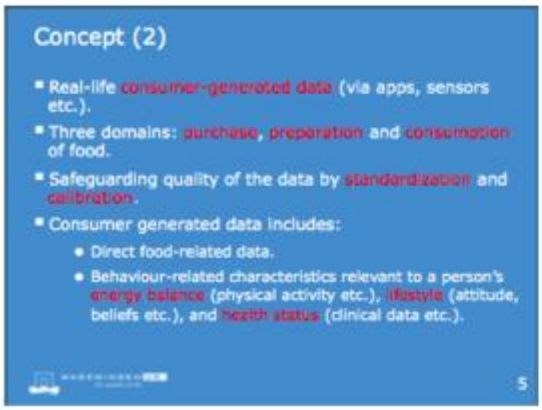
## *A2.2 Workshop presentations*

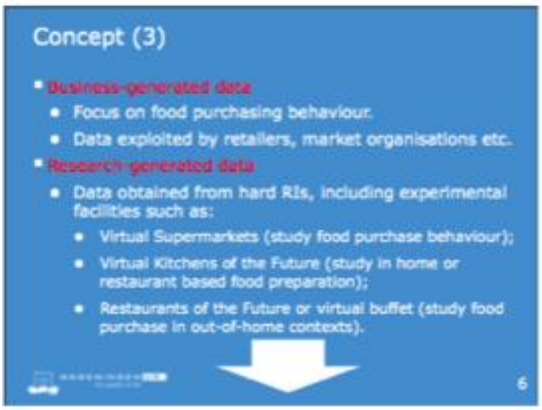
A2.2.1 10:30 – 10:35 Welcome (Siân Astley, EuroFIR AISBL - BE)

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

A2.2.2 10:35-10:45 Introduction (Paul Finglas, IFR - UK)








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

## 'Artist impression' of the RI consumer data platform

## Our approach of the design

- The RI Consumer Data Platform will be:
  - linking consumer (Big and Open) data: making connections between real-time and in-situ consumer data (phase 1) with business and research- generated consumer data from other RIs and from virtual labs (phase 2).
  - communicating analytical results, concepts and hypotheses (from those data).
- The design study will discuss and elaborate the requirements on: ICT regarding the data combining & data management, the business model, and the governance, legal (IPR etc.) an ethics issues (phase 3).

8

## The Work Plan

9

## Partners

- LEI-WUR (NL)
- DIL (DE)
- EuroFIR AISBL (BE)
- Jozef Stefan Inst. (SI)
- Wageningen Univ. (NL)
- Univ. of Surrey (UK)
- SP Techn. Research Institute (SE)
- Aalborg University (DK)
- Javier de la Cueva (ES)
- ETH Zurich (CH)
- EUFIC AISBL (BE)
- Institute of Food Res. (UK)
- CNR-ITIA (IT)
- CENS (FR)
- Aalto University (FI)
- GS1 Denmark (DK)

10

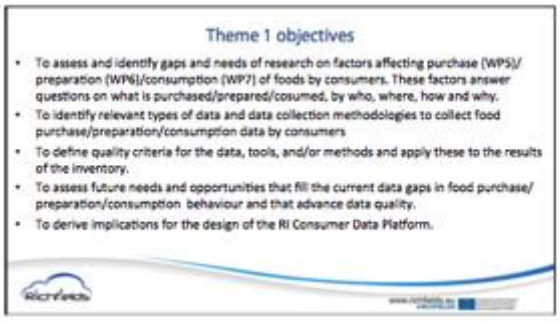
## Thank you for your attention!

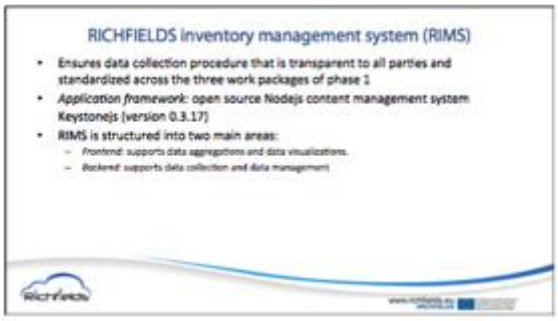
11



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

A2.2.3 10:45-11:05 Introduction: Phase 1 (Monique Raats, Uni Surrey – UK)

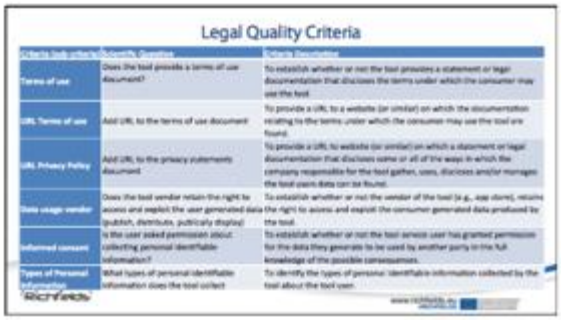


This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 654280.

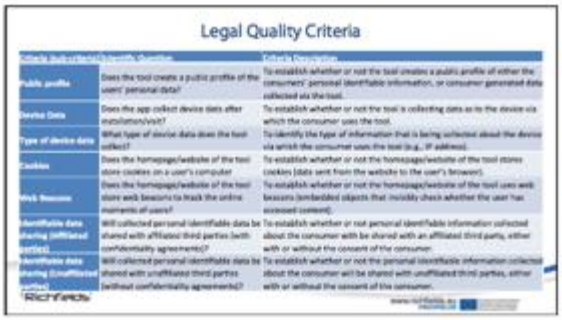
Technical Quality Criteria

| Criteria (sub-criteria) | Scientific Question | Criteria Description |
|---|---|---|
| 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). |
| Implements access protocols? | Can the data be accessed using a commonly used access protocol? | |
| 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. |
| 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. |



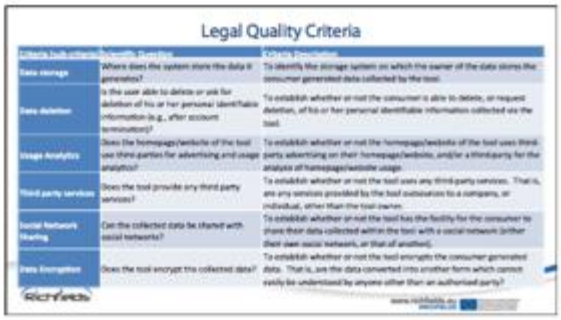
Legal Quality Criteria

| Criteria (sub-criteria) | Scientific Question | Criteria Description |
|---|---|---|
| Terms of use | Does the tool provide a 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. |
| 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. |
| 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. |
| 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. |
| 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. |
| 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. |



Legal Quality Criteria

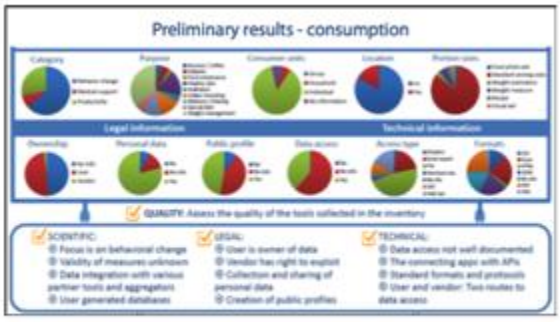
| Criteria (sub-criteria) | Scientific Question | Criteria Description |
|---|---|---|
| Public profile | Does the tool create a public profile of the users' personal data? | To establish whether or not the tool creates a public profile of either the consumers' personal identifiable information, or consumer generated data collected via the tool. |
| 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. |
| 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). |
| 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). |
| 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). |
| Identifiable data sharing (Affiliated Entities) | 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. |
| Identifiable data sharing (Unaffiliated Entities) | 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. |

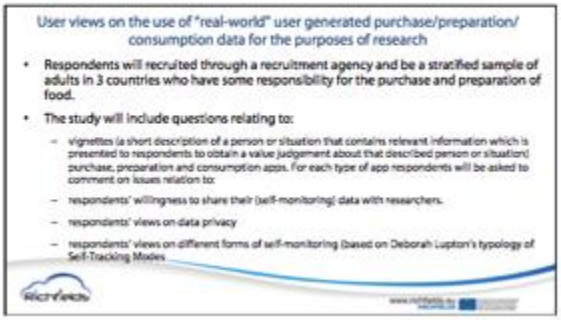


Legal Quality Criteria

| Criteria (sub-criteria) | Scientific Question | Criteria Description |
|---|---|---|
| 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. |
| 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. |
| 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 |
| 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. |
| 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). |
| 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 authorised party? |

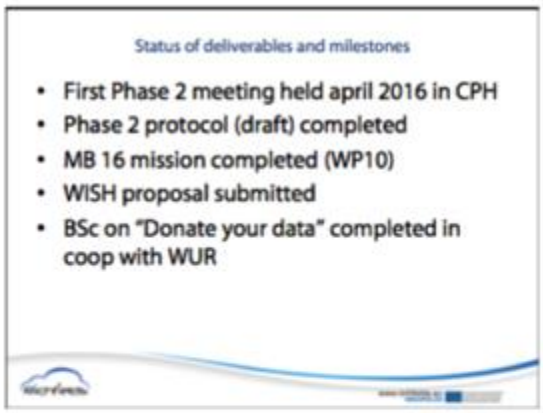
This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 654280.

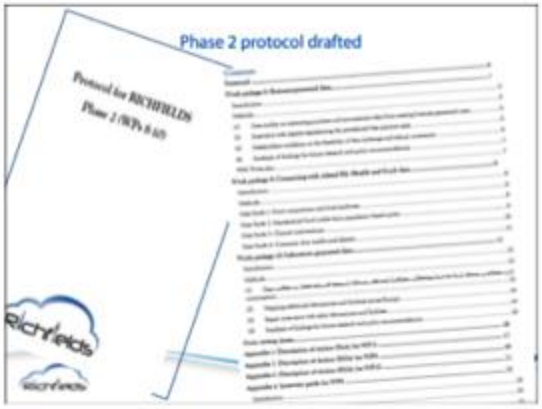


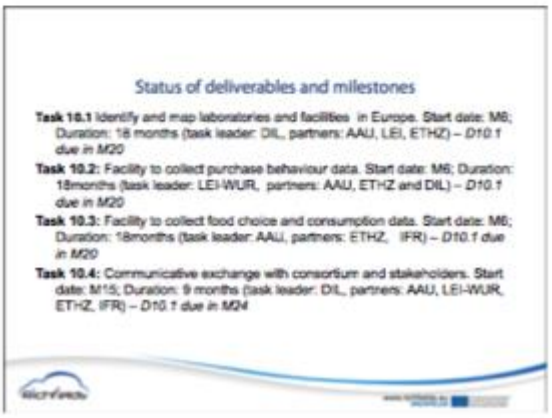
This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 654280.

A2.2.4 10:45-11:05 Introduction: Phase 2 (Bent Egberg Mikkelsen, Aalborg Uni – DK)



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






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

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

Food | Consumer | Health

## Phase 2

Bent Egberg Mikkelsen, (AAU)

www.richfields.eu
#RICHFIELDS

## Phase 2

WP 8 Bizz generated

WP 9 Related RIs (Health and Food) data

WP 10. Lab generated data (FSL, RoF, FFB)

## WP 8

| | DK | SE |
|---|---|---|
| Purchase (B2C) | GfK | ICA/Axfood |
| Procurement (B2G) | Statistics DK | City of Gothenburg |
| | | |

Protocol for RICHFIELDS
Phase 2 (WPs 8-10)

Data Capture and Management
Ethic and Governance
Business Model

## GfK & Statistics DK Sample

Data format and Web questionnaire

GfK

Statistic DK

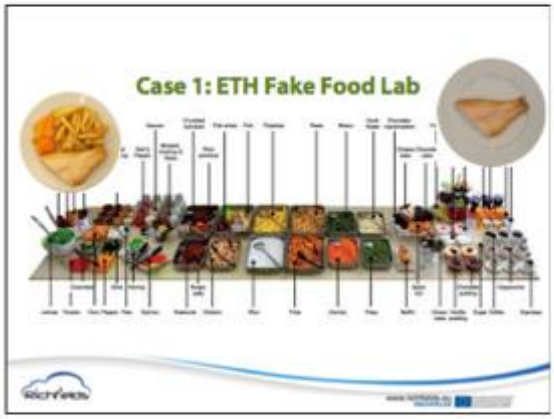
## Example –

ICA loyalty card data for one member

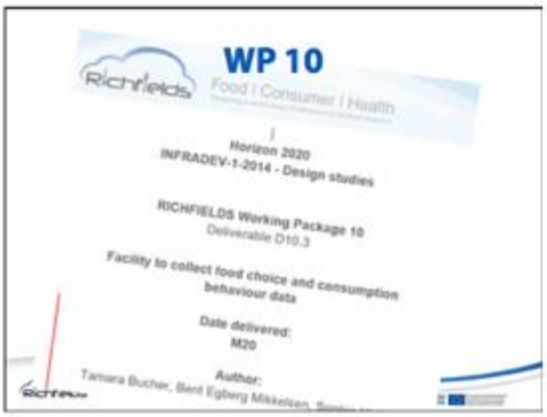
Personal information:
- Name
- Surname
- Personal number (contains information on gender and age)
- Customer number
- Address
- Membership date
- Phone number
- Email
- Type of household
- Loyalty card number



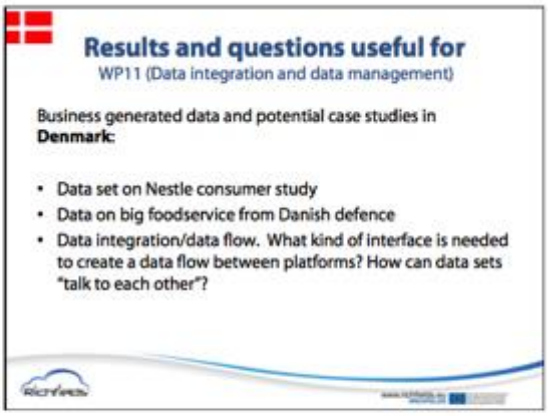
This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 654280.

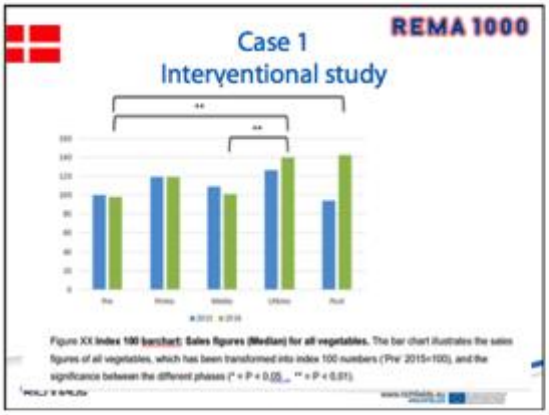


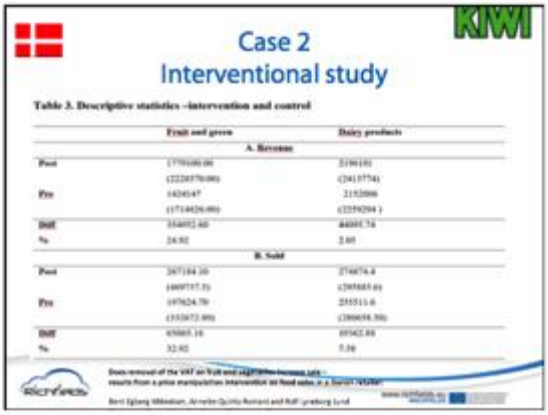
This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 654280.

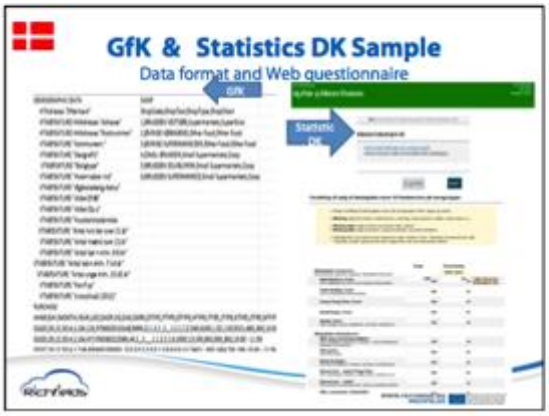


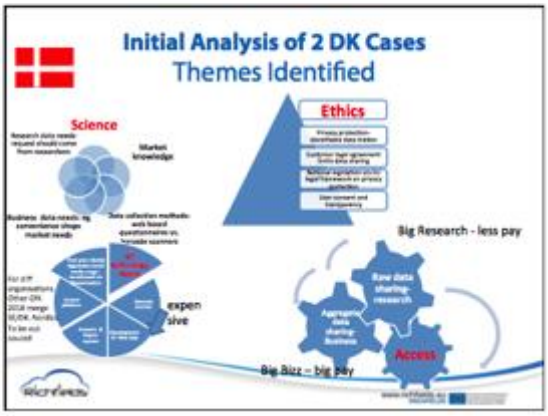
This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 654280.





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

## Results and questions useful for
WP11 (Data integration and data management)

Business generated data and potential case studies in **Sweden**:

- Loyalty card data from retailers (Ica, **Axfood**)
- Sales data (Statistics Sweden)
- Public procurement data (**Gothenburg Municipality**)
- Data from online retailers (mathem.se)
- Data from marketing firms (Adobe Marketing Cloud)

## Example –
ICA loyalty card data for one member

Personal information:

- Name
- Surname
- Personal number (contains information on gender and age)
- Customer number
- Address
- Membership date
- Phone number
- Email
- Type of household
- Loyalty card number

## Example –
ICA loyalty card data for one member

## Example –
Procurement data for Gothenburg Municipality

- Sales data from supplier collected twice a year
- One data sheet for each supplier
- Inconsistent structure and quality

## Results and questions useful for
WP12 (Business model)

- Loyalty card data is extremely important to retailers. The head of CRM has responsibility over data, but the decision to allow data to be used for scientific purposes comes from top management (Axfood)
- How do we approach businesses?
- What value does the Richfields project and a future RI provide for businesses?
- What are the alternatives to extracting business generated consumer data (bypassing retailers)?
- Business models. What kind of business ideas can be built on top of the data? What kind of consumer demands can be catered for, using the data commercially?
- Data quality. Third party auditing can add fidelity and credibility to data i.e using BC Tech
- Descriptions of Biases, weaknesses, sampling errors etc. needs to go hand in hand with datasets
- .

## Results and questions useful for
WP13 (Governance, Ethics and Final Design)

- Handling of personal information is regulated in Sweden by the PUL-law
- How does privacy regulation differ in European countries and how does this affect data extraction?
- People are becoming increasingly aware of personal data being collected by businesses and the value of that data. How will this affect future regulation? (rules for ownership of data)
- Ethics governance. It as proposed to distinguish between commercial applications and "a common cause"/the common good"/research/"become co-researcher"
- Access strategy. Partnership/NGO/PINGO approaches assumed a usefull strategy.?



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

## Status of deliverables and milestones

- Data collection 2' completed
  - DK
    Statistics (Informant: Martin)
  - GfK (Informant: Kenneth)
  - SE
  - ICA Axfood
  - Gothenburg Munic.
- Transcriptions * 2 made
- Samples of data structure provided (see examples)
- SE (SP)/ DK (AAU) agreed on common template analysis (CTA)
- See CTA sample
- Initial analysis carried out
- First themes identified (se specific slide)

## First draft of the agenda for Oct 6 meeting

- Selection (case 1 and 2)
- Groups: (Ethic, ICT & Business)
- Re-Interview. Interview nominated by group
- Interviews 15 mins
- Open meeting after lunch
- Panel discussion

## Scientific approach and publication strategy

- Phase
- Outreach to popular media initiated to create awareness on Big Data
- Webplatform on Big Food Data initated on www.capfoods.aau.dk *
- Open event on Big Food Data to be held on October 6 (tbc)

Richfields

www.richfields.eu
#RICHFIELDS

Partners:



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

## *A2.3 Workshop materials*

### A2.3.1 Summary of the B2R - GfK Denmark

GfK is market research organization conducting panel household survey on grocery purchase, analyses the data and provide market knowledge to different retail businesses in Denmark. The market knowledge helps the business to understand customers purchasing and provide services or products that could reflect in sales increase. GfK offers 3 different options for the house panellist to capture food purchasing data, pen and paper, digital scanner and the apps methods. Consumers can choose their preferred method. With the pen and paper, household panellist receives a diary each quarter and fills out their weekly purchases. For the barcode method, GfK provides free digital scanner to each panellist. The scanner is used to scan the barcodes of grocery purchased. The apps scanner, the most new and sophisticated can be downloaded onto a smart phone. The app is used to scans purchased food products. In all the methods panellist have to uploaded data onto the organizations' database website. The barcode information is supplemented by household panellist inputting additional information onto website in relation to when the product was purchased, price and the quantity.

**Business Generated Data**

The GfK household panel data structure is built on four pillars:

- Shop Demographics: includes information about shop chain code, channel type, and area location
- Household Demographics: includes information on age, income, family size and area location,
- Article (Product Characteristics): identifies what is in the shops, the barcode for a product and description, type of product.
- Movements (Purchases): depicts actual purchases of the households.
-

All these four are put together as a household food panel for the analysis of the purchasing behaviour

**User Needs**

The statistics are used by interest group retail sector and to understand consumer behaviour patterns.


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

**Sample of dataset format**

| | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| **Household Demographics** | | | | | | | | | | | |
| **Name** | Address | Tel | Email | Housing Type | Gender | Marital Status | Family Size | | | | |
| | | | | | | | | | | | |
| **Shop Demographics** | | | | | | | | | | | |
| **ShopCode** | ShopText | ShopType | ShopChain | | | | | | | | |
| | | | | | | | | | | | |
| **Articles** | | | | | | | | | | | |
| **Productgroup** | Dtype | Ftype | Gtype | Itype | Jtype | Package size | | | | | |
| | | | | | | | | | | | |
| **Movements (Purchases)** | | | | | | | | | | | |
| **HHNR** | Time | Day | Month | Year | Week | shop | EAN | Type | Total | Volume | Weight |


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

**Additional Capture Needs**

The current data does not cover out of home purchased and consumed food items by individual household members. These purchases are made usually in convenient shops and products are not scanned, creating gap in the completeness of the household purchasing data captured. The GfK respondent indicates that the industry is demanding food purchase data from the convenience market, which is currently not captured in the household panel survey. Capturing of this data will provide a complete picture of the total market.

**Centralized Data Platform**

Aside the Denmark office, GfK has branches and operates in the Scandinavian and other EU states with each having its own local platform to host house panel data and other related functions. Currently, the localized platform is being merged to become a centralized resource. There is no doubt that this will bring numerous benefits the entire organization. The respondent with experience in using the Danish platform could assert of it efficiency. This is because it is developed locally to provide solutions to the Danish operational needs and hoping that centralized platform could offer the same or better service.

**Data quality management**

Quality control is performed on the extracted households purchasing captured data. The demographic information, the size of the household, where they live is compared against national statistics to adjust for any discrepancy. For instance GfK found showing that the captured data contain a high proportion of elderly people than the national statistics. This is because the elderly are often willing to participate in household panels. In such cases measures were taken to adjust the data, "weigh" a little bit down for the household of the elder against the households that are young are not willing to participate. This is done internally before it is loaded to the local reporting solution that consultants and our customers are using.

**Ethnic and data governance**

***Piracy protection*** According the Gfk data acquisition from the clients is done in a transparent manner as they ensure that all the selected household panelist give consent of participation. The organization guarantees the household panelist that their personal information provided will be protected in compliance to the Danish law. The GfK representative revealed that the legal restrictions on personal data disclosure is strictly complied by the organization and blind any form identifiable data making available for use by businesses and researchers.

The organization does not disclose the full details of household panellist to research institutions because it is prohibited by the law. The institutions are provided with raw data that is not possible to identify individual panellist. If there is a need to provide details of household information the address, names, and other personal information, you have to inform, request has to be made to every household that participate to give their consent and they have to accept.

***Provider transparency*** Household panellist engaged in the survey is provided with detailed information in relation to the use of the data provided. The organization has a commercial motive

for collecting household purchasing data for retail business to improve operations and increase market share. The use of the data does not conflict with the objective of the households donating their purchasing data. The organization ensures that any transaction with the retail business does compromise the relational integrity with the household panellist.

***User transparency*** The household panellist is customers of various retail businesses that patronize Gfk services. These retailers are in constant competition to increase their customer share based on data driven market knowledge. GfK provides these solutions to counterparts, businesses, competing each for the customers capture. This raises transparency challenges for GfK as the household generated purchasing data on retail business client could be sold to other competitors and vice versa.

**Data sharing**

GfK analysed survey data and presents it in an aggregated format that makes  to explains household purchasing behaviour .This use of the aggregate data limits their clients from directly assessing household purchases data  and that individual house cannot be target for sales promotions. Currently JfK data are made available to researchers through sales and collaborative sharing. Research institutions such as universities without formal collaborative agreement on data sharing pay for the data any they get. There are also non-financial collaboration data sharing at the EU level, according to the respondents.

### A2.3.2 Summary of the G2R - Statistics Denmark

Statistics Denmark (SD) is a state-owned institution mandated to collect, compile, and publish data on sales of organic food and beverages in retail trade. The organic foods must meet Danish legislation on the marketing of organic products. SD has valid nationwide register of all businesses in Denmark. This means that SD uses extensive nationwide registers as a basis to form a representative sample for data collection. The data basis for the survey is reported from the major supermarket chains and wholesalers, who sell food to retailers. It is obligatory that businesses provide information requested by SD as part of the data collection. The annual data collection is carried out via a web form. The statistics are disseminated in News from Statistics Denmark and Statbank. Retail trade in food in Denmark is handled by a number of supermarket chains and department stores as well as a number of chains and independent grocers. The retail sector for organic foods is thus numerically quite extensive, but on the other hand also concentrated among a few players. This concentration means that it is possible to reduce the industry's overall reporting burden without this having a significant impact on the survey coverage and validity. The study population consists of two sub-populations. Partly supermarket chains and department stores as well as various wholesale chains. The supermarkets and department stores report actual sales of organic products, showing the net weight in kg and turnover in kroner, inclusive VAT. Wholesale chains corresponding report wholesale to retail sale showing the net weight in kg and turnover in kroner including VAT.

**Business Generator Data**

- Company demographics
- Product information
- Type of products sold
- Total weight of products sold
- Total volume of products sold
- Distribution of organic products by customer groups, e.g. Public institutions (hospitals, kindergarten-education, residential) Public and private work place canteens

**User Needs**

The statistics are used by interest groups, authorities etc. to assess the level and composition of organic consumption.

### A2.3.3 Paprika recipe manager

**TOOL TYPE**: APP
**PLATFORMS:** IOS, Android, Kindle, OS X
**LANGUAGES:** CS, DA, NL, EN, FI, FR, DE, HU, IT, JA, NB, PL, PT, ZH, ES, SV
**DOWNLOAD:** Google Play Store; Itunes App Store
**PRICE:** £3.99

Delightfully simple recipe management for everyone: from aspiring cooks to professional chefs. With web importing, grocery lists and meal planning, Paprika is the perfect kitchen companion. If you love to cook, Paprika will be most useful app you've ever downloaded! Use Paprika Cloud Sync to wirelessly sync between your devices and take your recipes with you everywhere you go! (Please note: each version of Paprika is sold separately for each de- vice, but Paprika Cloud Sync is a free service.) FEATURES • Create and modify your own recipes on any device. • Automatically download recipes from hundreds of websites with a single tap. • Integrated browser clipboard tools help you clip recipe information from any website. • Smart recipe view: cross off ingredients when you're done with them, and tap to highlight your current step. • Tools to help you cook: start multiple timers, pin active recipes to easily switch between them, and scale ingredients to your desired serving size. • Timers are automatically detected in your directions: simply tap on one to start! • Store nutritional information (automatically saved from sites which provide it) and personalized notes for each recipe. • Find your recipes in a flash with our powerful subcategory organization and advanced search tools which let you search by name, ingredient or source. • Intelligent grocery list automatically combines your ingredients and sorts them by aisle. • Use the pantry to keep track of what ingredients you already have. • Save time and money by planning your meals by the week and month! Create reus- able menus from your favorite meals. • Print and email support for recipes, grocery lists, and meal plans. Support for multiple recipe print sizes including index cards. • Share recipes via email (that other Paprika users can automatically import). • Import existing recipes from popular desktop apps such as MacGourmet, YummySoup!, Master- Cook & Living Cookbook. • Export your meal plans to iCal and your grocery lists to Reminders. • Manual backup and restore keeps your recipe collection safe. • Offline access. All of your data is stored locally, no internet connection is required to view your recipes. • Paprika Cloud Sync. Keep your recipes, grocery lists, and meal plans seamlessly synced between all your devices! • Bookmarklet. Capture recipes from your web browser straight into your Paprika Cloud Sync account.

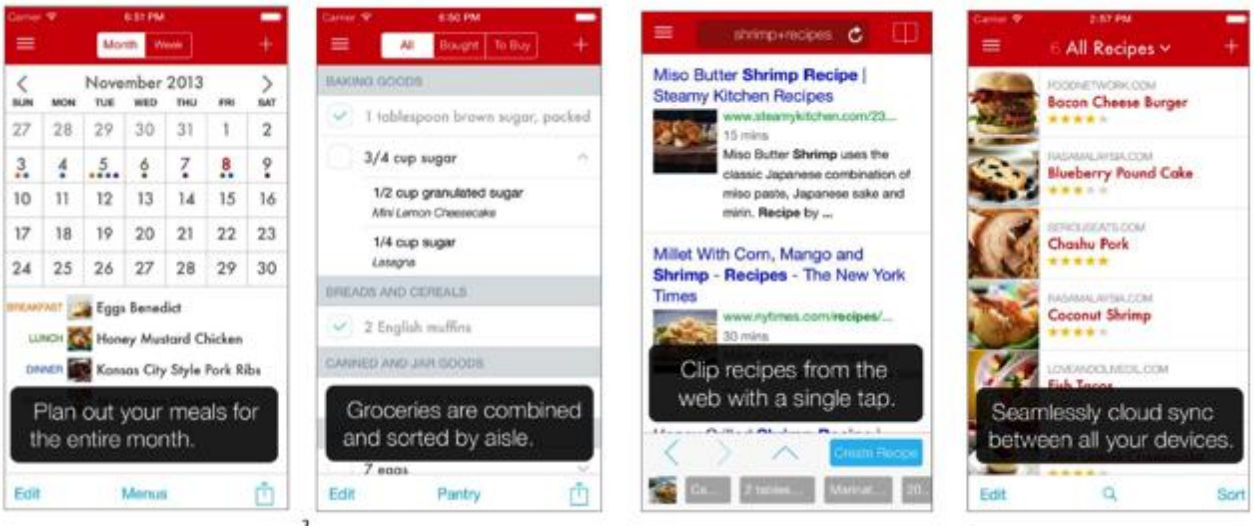
**USER INTERACTIONS**

- Create grocery lists, which can be automatically sorted by product category.
- Download recipes from websites and link to websites and other sources.
- Rate, 'favourite' and 'pin' recipes.
- Organise recipes, create user defined categories and subcategories.
- Link photographs to recipes.
- Create recipes within the app and modify existing recipes, including the addition of personal notes.

- Search for recipes in the database by ingredient or category/subcategory.
- Set timer for cooking and preparation time.
- Check off ingredients as you use them.
- Plan meals by week or by month.
- Email and print recipes and meal plans; sync with other apps (Calendar and Reminder).
- Interact with social media (Facebook and Twitter).

## USER GENERATED DATA

- Downloaded Recipes.
- Grocery list.
- User generated recipes and notes.
- List of favourites, pinned and rated recipes.
- Classification and grouping of recipes.
- User added photographs.
- Meal plans.
- Social media interactions.
- User information (email address, password, IP address).



### A2.3.4 Pantelligent

**TOOL TYPE**: APP; Sensor (frying pan)
**PLATFORMS:** IOS, Android, Watch IOS
**LANGUAGES:** EN
**DOWNLOAD:** Google Play Store; Itunes App Store
**PRICE:** FREE

This is the companion app for Pantelligent, the intelligent frying pan that helps you cook everything perfectly! Browse all of your favorite recipes, with each one backed by an exact time and temperature profile. Pantelligent gives you step-by-step directions and real-time feedback while you're cooking so you get it right the first time, and every time. Just pick a recipe and start cooking! Or you can set out on your own in Freestyle Mode, and get accu-rate temperature readings of the cooking surface, right where the food touches the pan. This lets you get a perfect golden-brown sear, sauté without ever burning, and simmer on low without over boiling. Don't have a Pantelligent? Get yours here: www.pantelligent.com Requirements: • iOS 7 or later and iPhone 4S or newer (including 4S, 5, 5C, 5S, 6, 6+, 6S and 6S+) • The Pantelligent frying pan also compatible with iPad 3rd gen or newer (including iPad Air and mini) and iPod 5th gen or newer.

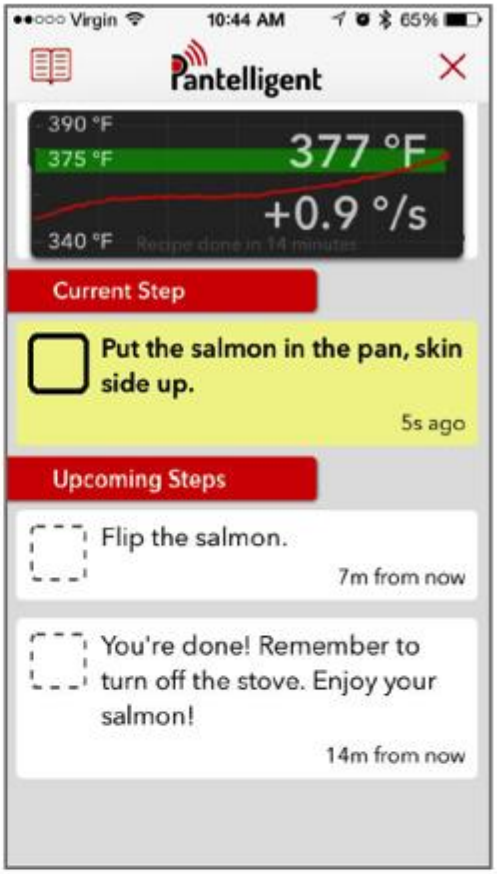
**USER INTERACTIONS**

- Browse recipe database
- Create recipes within the app
- Follow cooking instructions provided
- Enter characteristics about the food (e.g., thickness of food)
- Receive feedback in the form of notifications on cooking progress, time and temperature
- Read in app cooking tutorials

**USER GENERATED DATA**

- Recipe and food choice (including characteristics of the food, such as thickness)
- Cooking times (including date, time and duration)
- Preference for temperature unit (FO or Co)
- User created recipes
- Cooking history
- User information (name, email, username, password, IP address)

Virgin
10:38 AM
63%
Pantelligent
Simple Salmon
Prep Time: 5 minutes
Cooking Time: 19 minutes
Ingredients
1-2 salmon fillets
salt and pepper
1 tsp vegetable oil
¼ Tbsp butter


Virgin
10:44 AM
65%
Pantelligent
390 °F
375 °F
377 °F
+0.9 °/s
340 °F
Current Step
Put the salmon in the pan, skin side up.
5s ago
Upcoming Steps
Flip the salmon.
7m from now
You're done! Remember to turn off the stove. Enjoy your salmon!
14m from now

Virgin
10:38 AM
64%
Pantelligent
Simple Salmon
Mushroom Risotto

*A2.3.5 OCADO*

**TOOL TYPE**: APP; Sensor (frying pan)
**PLATFORMS:** IOS, Android, Kindle, Watch IOS
**LANGUAGES:** EN
**DOWNLOAD:** Google Play Store; Itunes App Store
**PRICE:** FREE

Order groceries with our app – it's a supermarket at your fingertips. Shop online, offline, anywhere, anytime on your iPad, iPhone or Apple Watch. • Save money with thousands of offers on your iPhone or iPad, and personalised Glance offers on your Apple Watch. • Order the freshest groceries and have everything delivered to your kitchen table in a one-hour slot that suits you. • Check your order status on any device to get up-to-the-minute info on your delivery. • As part of our Low Price Promise, we check your comparable shopping against Tesco's listed prices online. If we're not cheaper, we'll automatically send you a voucher for more than the difference – right up to £10. We're a British company delivering to England and Wales (for now...) We're expanding all the time; register on our app to check we deliver to your postcode. For more information about Low Price Promise please see ocado.com/LPPterms

**USER INTERACTIONS**

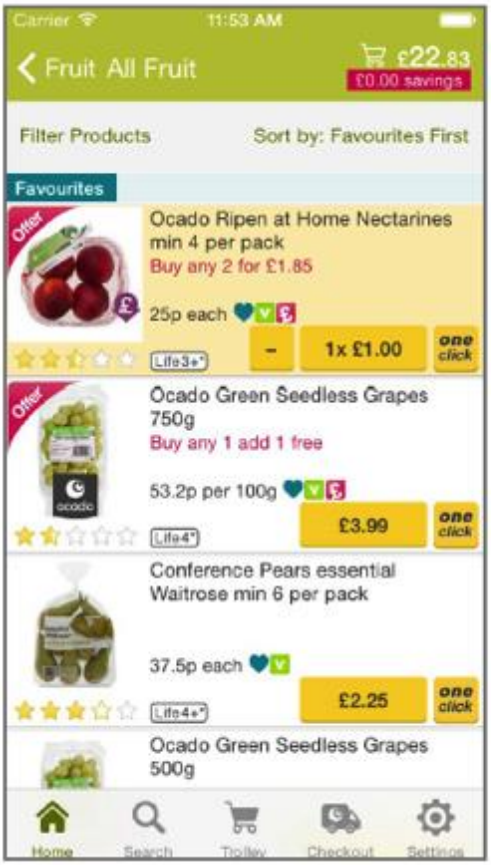
- Search for specific products or browse items by product category.
- Read consumer generated reviews and ratings (star system), and generate both ratings and reviews about individual products.
- Add items to shopping trolley.
- Look at app generated personalised offers.
- Purchase products through the app.
- Arrange delivery of products to a user determined postal address.
- Save lists of previously generated items for future use.

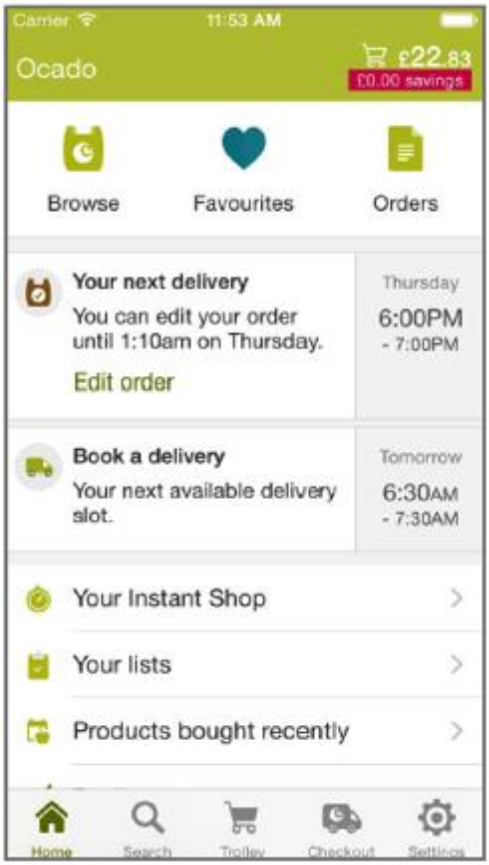
**USER GENERATED DATA**

- Product search history (a text list of user generated search terms).
- Shopping list (a list of products selected by the user, including quantity selected and information about these products as defined by the retailer e.g., product weight, nutritional information).
- Purchase history (a record of previous purchases made by the user).
- Information about the user (Including full name, address, postcode, phone number, username, password, financial information, IP address).

Carrier
11:53 AM
Trolley
Min. order £40
£0.00 savings
£22.83
12 items
Add to order
Checkout
Cupboard
Offer
Ocado Fairtrade Ripen at Home Bananas 7 per pack
£1.40
Buy any 2 for £2.50
Remove
Freezer
Home
Search
Trolley
Checkout
Settings

Back Ratings & Reviews
Ocado Lean Beef Steak Mince 5% Fat 500g
106 ratings
93 of 106 would recommend this product to a friend
Ratings snapshot
5 stars 84
4 stars 11
3 stars 2
2 stars 3
1 star 6
Add a rating or write a review
Fantastic quality
This is one of the best quality minced beef I've purchased at a supermarket. No tough, chewy bits and almost no fat which helps make a great lean meal.
Good value quality mince.
I found this mince was really good value and a lot nicer than the Waitrose mince and mince you find in other supermarkets. We used the mince for a chilli and really enjoyed it. Would purchase again.
Lean and lovely mince
I have bought this mince a few times...lean, very little fat comes out of it when cooked. Great price when on offer


Carrier
11:53 AM
Fruit All Fruit
£22.83
£0.00 savings
Filter Products
Sort by: Favourites First
Favourites
Offer
Ocado Ripen at Home Nectarines min 4 per pack
Buy any 2 for £1.85
25p each
1x £1.00
one click
Ocado Green Seedless Grapes 750g
Buy any 1 add 1 free
53.2p per 100g
£3.99
Conference Pears essential Waitrose min 6 per pack
37.5p each
£2.25
Ocado Green Seedless Grapes 500g
Home
Search
Trolley
Checkout
Settings


Carrier
11:53 AM
Ocado
£22.83
£0.00 savings
Browse
Favourites
Orders
Your next delivery
You can edit your order until 1:10am on Thursday.
Edit order
Thursday
6:00PM
- 7:00PM
Book a delivery
Your next available delivery slot.
Tomorrow
6:30AM
- 7:30AM
Your Instant Shop
Your lists
Products bought recently
Home
Search
Trolley
Checkout
Settings

*A2.3.6 MyfitnessPal*

**TOOL TYPE**: APP; Website
**PLATFORMS:** IOS, Android, Windows Phone
**LANGUAGES:** DA, NL, EN, FR, DE, ID, IT, JA, KO, MS, NB, PL, PT, RU, ZH, ES, SV, ZH, TR
**DOWNLOAD:** Google Play Store; Itunes App Store
**PRICE:** FREE

Lose weight with MyFitnessPal, the fastest and easiest-to-use calorie counter for iOS. With the largest food database by far (over 5,000,000 foods) and amazingly fast and easy food and exercise entry, we'll help you take those extra pounds off! And it's FREE! There is no better diet app - period. • Consumer Reports #1 rated diet • PC Magazine Editor's Choice Selection • #1 Health and Fitness app for 4 years straight Also featured in the NY Times, Wall Street Journal, Wired, USA Today, Family Circle, Marie Claire, NBC, CNET, Shape, the Today Show and more. USER REVIEWS • "The first diet tool that has ever worked for me! I've lost 30 pounds!" • "This is the best calorie counter, free or paid, and I've tried them all." • "The food database is HUGE! I've NEVER had a missing food." • "This app takes seconds to use, it's that simple." WHY WE'RE BETTER THAN OTHER APPS • FREE – sign up with no strings attached • Track your diet and exercise in less than 5 minutes a day! It's that fast and easy • Largest food database of any iPhone calorie counter – over 5,000,000 foods and growing daily • Easiest and fastest food entry – remembers your favorites, add multiple foods at once, save and add entire meals, and more. THERE IS NO FASTER OR EASIER APP THAN MYFITNESSPAL. • Connect with over 50 devices and apps including Apple Health, Fitbit, Jawbone UP, Garmin, MapMyFitness, Runkeep-er, Strava, Runtastic, Misfit, Withings, and more! • Recipe importer - cook a lot? Our recipe importer lets you visit any recipe on the web and easily import and track it with just a tap! It feels like magic! • Built-in step tracker - iPhone 5S/6/6+ users can track steps right from their phone, no separate tracker required • MYFITNESSPAL WORKS – our members have lost almost 200 MILLION POUNDS combined! EASIEST APP TO TRACK DIET & EXERCISE • 5,000,000+ food database of global items and cuisines. Virtually every food you eat is in our database al-ready • Barcode scanner - track a food just by scanning it's barcode. Over 4 million barcodes recognized! • Step tracker - iPhone 5S and iPh-one 6/6+ users can track steps and overall calorie burn right from their phone. No separate tracker required! • Track all major nutrients: calories, fat, protein, carbs, sugar, fiber, cholesterol, and more CONNECT APPS & DEVICES • Easily connects and seamless integrates with over 50 apps and devices including Apple Health, Fitbit, Jawbone UP, Garmin, MapMyFitness, Runkeeper, Strava, Runtastic, Misfit, Withings, and more! Works with virtually every fitness app and device. GET SUPPORT • Connect with friends and easily track and motivate each other, or meet new friends from the MyFitnessPal community! GOALS & REPORTS • Received personalized goals based on your individual diet profile, or enter your own goals if you've gotten specific recommendations from a doctor, nutritionist, etc. • Gain insights into where your calories and nutrients are coming from and how to make healthier choices • View charts of your progress over time for motivation Why wait? It is free to sign up so you have nothing to lose except those extra pounds! Download us and start leading a healthier life today! SUBSCRIPTION PRICING & TERMS MyFitnessPal is free to download and use. Should you choose to upgrade to MyFitnessPal Premium, we offer two auto-renewing subscription options: $9.99/month $49.99/year Payment will be charged to your credit card through

your iTunes account at confirmation of purchase. Subscription renews automatically unless cancelled at least 24 hours prior to the end of the subscription period. There is no increase in price when renewing. Subscriptions can be managed and auto-renewal turned off in Account Settings in iTunes after purchase. Once purchased, refunds will not be provided for any unused portion of the term. Read our full Terms of Service and our Privacy Policy at:
http://www.myfitnesspal.com/account/terms_and_privacy

### USER INTERACTIONS

- Input foods and beverages consumed; track diet (including calories, fat, protein, carbs, sugar, fibre, cholesterol).
- Search food database.
- Import recipes from other sources.
- Interaction with the MyFitnessPal Community and social media (including Facebook and Twitter).
- Set personal goals and targets, and receive personalised fitness goals and targets based on inputs.
- View personal progress through food diaries and charts.
- Sync with personal fitness device and other apps (e.g., Apple health, Fitbit, Jawbone, Garmin, MapMyRun, RunKeeper).

### USER GENERATED DATA

- Self-reported food diary data (including calories, fat, protein, carbs, sugar, fibre, cholesterol) and water drunk.
- Portion sizes and meal patterns (times of meals e.g., breakfast, lunch, dinner, snack).
- Search history (text list of search terms).
- Product barcode information.
- User information (location, date of birth, email, username, password, name, address).
- Height, weight, age.
- Steps / distances (depending on device interaction).
- Social media posts.
- Fitness goals and targets.

Quickly scan barcodes
to add nutrition facts.
Scan a Barcode

Sync automatically
with the apps you love.
Apps & Devices
Featured
All Apps

Our users have lost over
200 million pounds.
Blog
Michigan Mom Loses 100
Pounds with
MyFitnessPal

More than 6 million foods
in our database.
Add Food

Keep track of
what you're eating.
Edit
Diary
Today

A2.3.7 Lloyds bank mobile banking

**TOOL TYPE**: APP
**PLATFORMS:** IOS, Android, Windows Phone
**LANGUAGES:** EN
**DOWNLOAD:** Google Play Store; Itunes App Store
**PRICE:** FREE

Bank the way that suits you, with the Lloyds Bank mobile banking app. Using our app is easy and very secure. You can: - Log on with just 3 characters - Check your balances, transactions and pending payments - Make transfers, payments and standing orders - Pay your credit card - Pay a contact and make international payments - Earn as you spend with Everyday Offers - Apply for loans, savings, cards and more - Report lost or stolen cards and order re-placements - Tell us if you're planning to travel abroad - Reset your Internet Banking password - Update the phone number and email we have for you - Find a Lloyds Bank branch, Cashpoint® or LINK ATM GETTING STARTED New to this app? You'll need to set it up by registering your device first. You'll need: - An up-to-date phone number regis-tered with us - Lloyds Bank UK personal, Islands personal or sterling international account - Internet Banking User ID, password and memorable information KEEPING YOU SAFE ONLINE We use the latest online security measures to protect your money, your personal information and your privacy. As a Lloyds Bank Mobile Banking customer, you automatically benefit from our online and mobile banking guarantee. You can find further information about our fraud guarantee by checking the Security FAQs in this app or by visiting http://www.lloydsbank.com/security.asp We know security is important to you and we want you to feel safe online. When you receive emails from us we will always greet you personally using your title and surname and include either the last 4 digits of your account number or the last part of your postcode '*** 1AB'. For text alerts, check they come from LLOYDSBANK. IM-PORTANT INFORMATION Mobile Banking is available to our UK personal (and some Channel Islands, Isle of Man and WorldWide Service) Internet Banking customers. Services may be affected by phone signal and functionality. Terms and conditions apply. Everyday Offers is available to Lloyds Bank UK personal current account customers with a debit/credit card aged 18+ who bank online. Terms and conditions apply. Device and card restrictions apply to Apple Pay. When you use this app we collect non-personally identifiable location data for security purposes, to help combat fraud,fix bugs and improve the services we can provide in the future. In the UK the Branch/ATM finder uses location based services. We, LINK Scheme, Google and/or Apple may access data about your location (i.e. GPS signals from your mobile device) which can be used to approximate your location (such as your mobile ID). Cash-point® is a registered trademark of Lloyds Bank plc. Lloyds Bank plc (registered in England and Wales (no. 2065), registered office: 25 Gresham Street, London EC2V 7HN). Authorised by the Prudential Regulation Authority and regulated by the Financial Conduct Authority and the Prudential Regulation Authority under registration number 119278. Eligible deposits with us are protected by the Financial Services Compensation Scheme (FSCS). We are cov-ered by the Financial Ombudsman Service (FOS). Please note that due to FSCS and FOS eligibility criteria not all Lloyds Bank customers will be covered. This app is intended for UK residents unless otherwise stated.

Holders of accounts in Jersey, the Bailiwick of Guernsey and the Isle of Man can find compensation scheme details and other relevant information at http://international.lloydsbank.com/

## USER INTERACTIONS

- Monitor bank balance, transactions and payments.
- Make payments and transfer money.
- Set up standing orders.
- Pay credit card bill.
- Apply for loans, savings products and bankcards.
- Report lost or stolen cards.
- Reset password and update user information (including address, email and phone number).
- Locate an ATM or physical branch.

## USER GENERATED DATA

- Banking profile (including bank balance, income, loans, credit history).
- Spending profile (time, date, retailers).
- User profile (name, address, phone numbers, user ID, password, security questions, IP address).
- GPS data.

Your Memorable Information
Please enter the indicated characters from your memorable information
2nd
4th
8th
Log on quickly and securely with just your Memorable Information.

Menu
Make a Transfer
Pay an Account
Pay a Contact
From
Classic Account
£87.56
12-34-56 | 12345678
Available £87.56
To John Doe
£20.00
Ref: Fridaylunch
Review payment
Pay a contact, an account or transfer funds between your eligible accounts.

Menu
Classic Account
12-34-56 | 12345678
Direct Debits
Standing orders
Set up a standing order
John Doe
£2.00 / Monthly
Next due: 25 Oct 14
Expiry date: Until further notice
Ref: Monthlytransfer
Paid to: 12-34-56 | 87654321
Delete
Amend
Jane Stevens
£259.80 / Quarterly
Next due: 30 Oct 14
Expiry date: Until further notice
Manage your Direct Debits and standing orders on the go - with our secure app.

*A2.3.8 Fitbit*

**TOOL TYPE**: APP, Website, Wearable Device
**PLATFORMS:** IOS, Android
**LANGUAGES:** EN, FR, DE, IT, JA, KO, ZH, ES
**DOWNLOAD:** Google Play Store; Itunes App Store
**PRICE:** FREE

Live a healthier, more active life with Fitbit, the world's leading app for tracking all-day activity, workouts, sleep and more. Use the app on its own to track basic activity and runs on your phone, or connect with one of Fitbit's many activity trackers and the Aria Wi-Fi Smart Scale to get a complete picture of your health—including steps, distance, calories burned, sleep, weight, and more. TRACK ACTIVITY: Accurately record your steps and distance with Mo-bileTrack when you carry your phone. For all-day tracking of stats like calories burned, active minutes, and sleep, pair the app with a Fitbit tracker. RUN SMARTER: Enhance runs, walks and hikes by using MobileRun to track your pace, time and distance. You can also control your music, get voice cues and use your phone's GPS to map your routes. (Continued use of GPS running in the background can dramatically decrease battery life.) RECORD WORKOUTS: Use your Fitbit tracker to track your exercise, then check the app to see your stats, their impact on your day, and how your performance is improving. MONITOR HEART RATE: Use a Fitbit tracker with PurePulse™ to analyze heart rate graphs in the app. Identify trends, manage stress and see the results of your workouts. Review resting heart rate trends to see when your fitness is improving. LOG FOOD FASTER: Easily log calories with our bar-code scanner, calorie estimator, and expanded food database of more than 350,000 foods. See your meal history at a glance, and get nutritional insights. MEASURE HYDRATION: Quickly log your water intake to make sure you're properly hydrated during workouts and throughout the day. SET & MANAGE GOALS: Create weight, nutrition and exercise goals, and start a food plan to stay on track. Then get a visual picture of your progress with colorful, easy-to-read charts and graphs. SEE HOW YOU SLEEP: Set sleep goals in the app, and use a Fitbit tracker to monitor how much time you spent awake, restless or peacefully sleeping. SHARE & COMPETE: Connect with friends and family by sharing stats, sending direct-messages, and competing on the leaderboard or in Fitbit Challenges. STAY MOTIVATED: Get a nudge in the right direction with notifications that pop up when you're close to reaching a goal or have already met one. SYNC WIRELESSLY: Fitbit trackers sync your stats to computers and 200+ leading devices so you can continuously track your progress without needing to plug in. MANAGE WEIGHT: Connect wirelessly to the Aria Wi-Fi Smart Scale to seamlessly track your weight, BMI, lean mass and body fat percentages, and to see your weight trends over time. Learn more about Fitbit products and services at www.fitbit.com.

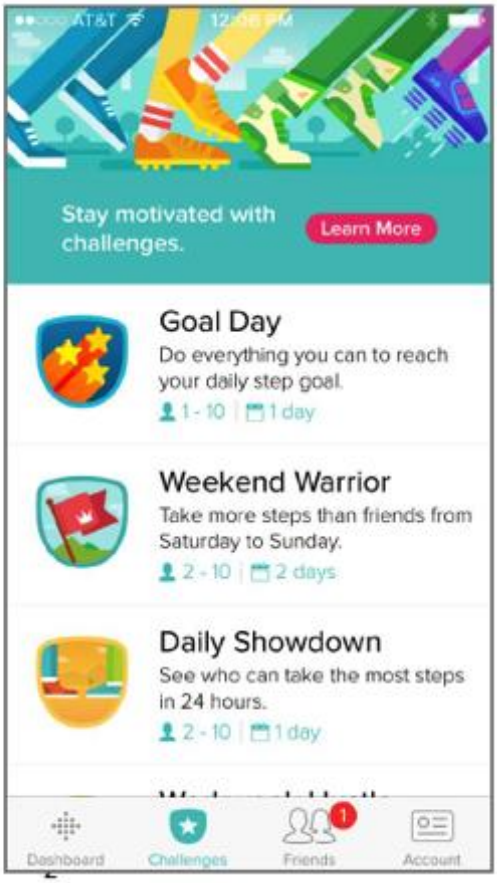
**USER INTERACTIONS**

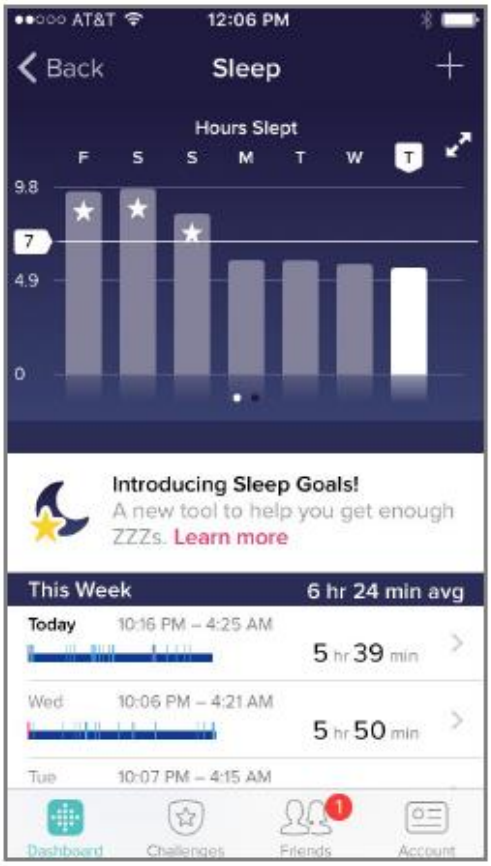
- Track activities (including, sleep, distance walked/ran, calories burned, sleep, weight)
- Connect with wearable device.
- Receive feedback on activity level by viewing graphs and personal statistics.
- Map routes walked and ran.
- Receive notifications and feedback in the form of voice commands.

- Record food intake—including water— in a food diary.
- Search database of foods and scan product barcodes for nutrition information.
- Set personal goals and targets for weight, nutrition and exercise.
- Interact with social media.

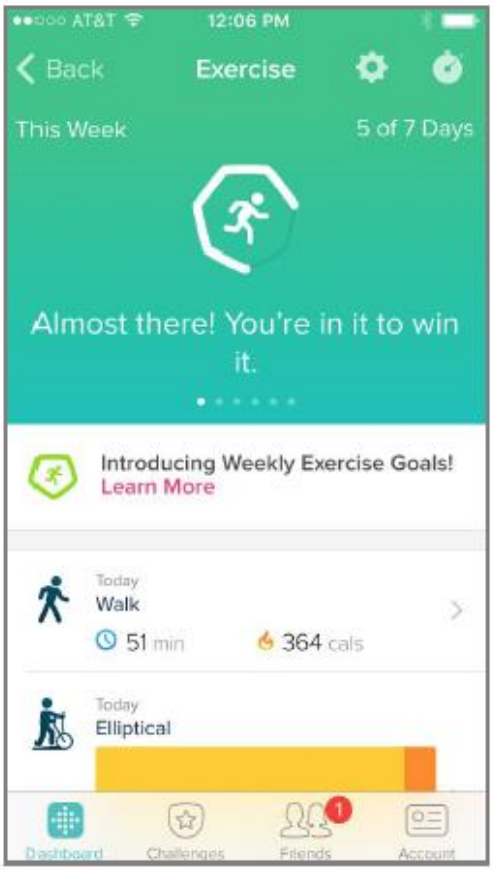
### USER GENERATED DATA

- Data on number of steps taken /distance, duration of sleep, and frequency and type of exercise.
- Personal goals and targets (weight, nutrition, fitness, sleep).
- Self-reported food diary data (calories, nutrition information) and water drunk.
- Product barcode information and food database search history.
- User information (location, date of birth, email, username, password, name, address, IP address).
- Height, weight, age.
- Steps / distances (depending on device interaction).
- Social media interactions.
- Fitness goals and targets.
- Notifications sent to user.




AT&T
12:06 PM
Back
Sleep
Hours Slept
F S S M T W T
9.8
7
4.9
0
Introducing Sleep Goals!
A new tool to help you get enough ZZZs. Learn more
This Week
6 hr 24 min avg
Today 10:16 PM – 4:25 AM
5 hr 39 min
Wed 10:06 PM – 4:21 AM
5 hr 50 min
Tue 10:07 PM – 4:15 AM
Dashboard
Challenges
Friends
Account

AT&T
12:55 PM
Back
Weight
Weight (lbs)
164
159
155
11/3
Today
GOAL PROGRESS
Edit Goal
0.5 lbs
lost since 12/1
5.8 lbs to go and this goal is in the bag.
Log Weight
This Week
161.1 lbs avg
Today
160.8 lbs
16.7 % Fat
Dashboard
Challenges
Friends
Account


AT&T
12:06 PM
Back
Exercise
This Week
5 of 7 Days
Almost there! You're in it to win it.
Introducing Weekly Exercise Goals!
Learn More
Today
Walk
51 min
364 cals
Today
Elliptical
Dashboard
Challenges
Friends
Account

*A2.4 Sign-in sheets*

F. Sign-in sheet

| | Name | Organisation | Signature |
|---|---|---|---|
| 2?/ 1 | Aida Turrini | CREA- Alimenti e Nutrizione) (IT) | |
| 2 | Paolo Colombani | Independent consultant (CH) | |
| 1 | James Lay | Food Angels Ltd. (UK) | |
| 1 | Pieter Francois | University of Hertfordshire and University of Oxford (UK) | |
| 1 | Roel van der Heijden | University Medical Center Groningen (NL) | |
| 1 | Giulia Vilone | Creme Global Ltd. (IE) | |
| 2 | Maud Alligier | European Clinical Research Infrastructure Network (ECRIN) | |
| 2 | Paul Allington | The Code Guy (UK) | |
| 2 | Kristrún Gunnarsdóttir | Centre for Research in Social Simulation (CRESS), University of Surrey (UK) | |
| F2 | Monique M. Raats | Consumer Behaviour and Health Research Centre, University of Surrey (UK) | |
| 2 | Bent Egberg Mikkelsen | Aalborg University (DK) | |
| 1 | Paul Finglas | Institute of Food Research (UK) | |
| N2 2 | Siân Astley | EuroFIR AISBL | |
| 2 | Angelika Mantur | EuroFIR AISBL | |
| F1 | Lada Timotijevic | Food, Consumer Behaviour and Health Research Centre, University | |

www.richfields.eu
#RICHFIELDS

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

| | | | |
|---|---|---|---|
| | | of Surrey (UK) | |
| N1 | Kerry Ann Brown | University of Surrey (UK) | |
| 2 | Naomi Klepacz | Consumer Behaviour and Health Research Centre, University of Surrey (UK) | Naomi Klepacz |
| 1 | Charo Hodgkins | Consumer Behaviour and Health Research Centre, University of Surrey (UK) | |
| 1 | Marcus Maringer | Wageningen University (NL) | |
| 2 | Barbara Korousic Seljak | Institut Jozef Stefan (SI) | B. K. Seljak |

Javier de la Cruz Javier de la Cruz Helena Cruz



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

Richfields

Consumer data | Food | Health

RICHFIELDS is designing a world class research infrastructure providing data, tools or services to enable study of consumer food-related behaviour

**Registration Form**
**RICHFIELDS STAKEHOLDER WORKSHOP**
**27.09.2016, Schiphol, Netherlands**

| Name | Organisation | Signature |
|---|---|---|
| Alligier Maud | ECRIN | — |
| Allington Paul | The Code Guy Ltd | |
| Astley Sian | EuroFIR | — |
| Brown Kerry Ann | University of Surrey | |
| Colombani Paolo | Consulting Colombani | — |
| Finglas Paul | Institute of Food Research | |
| Francois Pieter | University of Hertfordshire | |
| Gunnarsdottir Kristrun | University of Surrey | — |
| Hodgkins Charo | University of Surrey | |
| Klepacz Naomi | University of Surrey | — |
| Korousic Seljak Barbara | Jožef Stefan Institute | — |
| Lay James | Food Angels | |
| Mantur Angelika | EuroFIR | |
| Maringer Marcus | Wageningen University | |
| Mikkelsen Bent Egberg | Aalborg University | — |
| Raats Monique | University of Surrey | — |
| Timotijevic Lada | University of Surrey | |
| Turrini Aida | CREA | |
| van der Heijden Roel | UMC Groningen | |
| Vilone Giulia | Creme Global | |

Javier de la Cueva Javierdlcueva

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

Richfields

Consumer data | Food | Health

RICHFIELDS is designing a world-class research infrastructure (providing data, tools or services) to enable study of consumer food-related behaviour

**Registration Form**
**RICHFIELDS STAKEHOLDER WORKSHOP**
**27.09.2016, Schiphol, Netherlands**

| Name | Organisation | Signature |
|---|---|---|
| Alligier Maud | ECRIN | |
| Allington Paul | The Code Guy Ltd | |
| Astley Sian | EuroFIR | |
| Brown Kerry Ann | University of Surrey | — |
| Colombani Paolo | Consulting Colombani | |
| Finglas Paul | Institute of Food Research | — |
| Francois Pieter | University of Hertfordshire | — |
| Gunnarsdottir Kristrun | University of Surrey | |
| Hodgkins Charo | University of Surrey | — |
| Klepacz Naomi | University of Surrey | Naomi Klepacz |
| Korousic Seljak Barbara | Jožef Stefan Institute | |
| Lay James | Food Angels | — |
| Mantur Angelika | EuroFIR | |
| Maringer Marcus | Wageningen University | — |
| Mikkelsen Bent Egberg | Aalborg University | |
| Raats Monique | University of Surrey | |
| Timotijevic Lada | University of Surrey | — |
| Turrini Aida | CREA | — |
| van der Heijden Roel | UMC Groningen | — |
| Vilone Giulia | Creme Global | — |

www.richfields.eu

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

#RICHFIELDS

## Annex 3: Workshop notes: presentation and breakout session

*A3.1 Group assignment and questions*

| Group 2 | Group 1 |
|---|---|
| Paolo Colombani (CH) | Aida Turrini (IT) |
| Maud Alligier (FR) | James Lay (UK) |
| Paul Allington (UK) | Pieter Francois (UK) |
| Kristrún Gunnarsdóttir (UK) | Roel van der Heijden (NL) |
| Bent Egberg Mikkelsen (DK) | Guilia Vilone (IE) |
| Angelika Mantur (BE) | Charo Hodgkins (UK) |
| Naomi Klepacz (UK) | Marcus Maringer (NL) |
| Barbara Koroušić Seljak (SI) | Javier de la Cueva (ES) |
| | Paul Finglas (UK) |
| | |
| **Rapporteur:** | **Facilitator:** |
| (Group 1) Kerry Brown (UK) | (Group 1) Lada Timotijevic (UK) |
| (Group 2) Siân Astley (BE) | (Group 2) Monique Raats (UK) |

**Breakout Session 1 Questions: Understanding the nature of the data**

- What do these data allow you to do within your discipline? What research questions can they help answer (or allows you to ask) within your discipline?
- What unit of observation does it represent – who or what can you attribute to these data to – the individual, household, an organization, geographical location, social interaction etc.?
- What are the shortcomings of the data?
- How useful would these data be to a (discipline) studying food-related phenomena?

**Breakout Session 2 Questions: Developing a meta-data**

- What do we need to know about data to make these useful for research of food related phenomena (e.g. intake, practices, health)?
- What do we need to be recording about the data in a meta-data document?

*A3.2 Measurement break-out session feedback*

**Breakout Session 1 (data):**

- What do these data allow you to do within your discipline? What research questions can they help answer (or allows you to ask) within your discipline?
- What unit of observation does it represent – who or what can you attribute to these data to – the individual, household, an organization, geographical location, social interaction etc.?
- What are the shortcomings of the data?
- How useful would these data be to a (discipline) studying food-related phenomena?

### *Group 1: Lada Timotijevic (UK), reported by Kerry A. Brown (UK)*

*Case studies discussed: Statistics Denmark (G2R) (grocery purchase at the household level), Paprika recipe manager (recipes), Fitbit (health and wellbeing)*

Discussion – variety of personal and professional experience comments:
CH// Issue of missing/inaccurate data - may or may not enter all this data, integrated with existing software, sporadic use?
JL// Exercise application can lead to competition of use – app an intervention in itself.
CH// Idea to capture information not construct a research design to capture information. Can potentially look at trends within a person, over time or potentially aggregate to a population level? [KAB comment// how to impose a research design on naturally occurring data http://methods.sagepub.com/video/what-is-naturally-occurring-data]
RvdH// Cohort study running in NL. Looking at the potential to add prospective data rather than rely on recall questionnaires.
LT// How might this be achieved? Can you define this cohort, is the information available on the cohort demographics?
RvdH// The biobank is starting to explore how to add real-time data. Cohort includes predominantly 40-50 year olds and their families. Can identify individuals who are male/female etc. then can potentially ask a certain group to participate in a study.
JdlC// Data at an individual or aggregated level – could one manipulate the cohort?
RvdH// Idea is to observe the cohort rather than intervene.
CH// What really is the difference between data from a population who was asked to wear a fitbit as part of a research study and those that choose to wear a fitbit of their own accord? Is one data set better than the other and why? [KAB comment// There are differences and these differences are the interesting ones for RICHFIELDS to spell out and attempt to question]
PieterF// How many of the cohort already wear a fitbit?
GV// A limitation of fitbits is that children won't be included in any data sample. An advantage is that this data can be useful for investigating long terms trends and correlations.
LT// What needs to be taken into account to interpret a trend?
GV// Demographics.
JL// Income level.
CH// Which groups might not be able to manage the use of these tools (such as fitbit), which groups might not be health motivated, which groups might not be able to afford such tools. All useful

information to identify if any data is or is not being gathered on traditionally hard to reach populations.
RvdH// Analysis can compare the results of “big data” with that of research led data and see if these are valid or not [KAB comment// how frame and decide what is and what is not valid is an interesting question from a disciplinary and science/non-science point of view].
LT// How can data be verified? Does it matter if it is or it isn’t verified?
CH// Data itself doesn’t present with enough information to make inferences.
JL// Fitbit users can become overly conscious about everything they do/fitbit produces about them. That is how one becomes committed to the gadget. Can be driven by the social groups they interact with, can represent pro-active individuals who are wishing to take responsibility for their own health.
JdlC// Social responsibility that comes with producing these tools – could they encourage negative thoughts and behaviours?
PieterF// Problem of a large dataset being presented which does not align with a RQ. Would it be possible to select a sub-section of the population/dataset available where a RQ could be investigated using the “big data” and some understanding of a group could be established. The question would then be to what extent could any knowledge be generalized to the wider population?
LT// How could a group/community be defined?
PieterF// This is an interesting RQ in itself, such who they are, is there is a shared value in the community, how can they be defined.
PaulF// More importantly can one even access the data to use it?
CH// Access the fitbit data to target a cohort?
MM// Every individual will have provided data that can be publically accessible (a public API). They might wish to change their privacy settings but some information will still be available. The other option is to go via the merchandiser and pay to access their aggregated data. Value is then added by linking data to other platforms e.g., facebook etc. The larger networks do this to enrich their datasets. RICHFIELDS could do the same thing.
AT// To assess a trend one needs to have a baseline. Information would be required from perhaps when an app was first downloaded.
LT// Information at an individual or aggregated level?
MM// It is possible to do both. It might not be ethical to do it at an individual level but currently it would be possible as this information is out there and publically available.
CH// Have to consider the issues of forced consent (tick terms and conditions to proceed) rather than informed consent. An RI would have to comply with ethics, so this is an issue that will need to be addressed (just because one can doesn’t mean that one should).
JdlC// Ethical questions include, who is the owner of the data (at the individual or aggregated level); what is the context/social contract (is it citizen science?); are we taking scientists out of the picture and asking the general public to be their own researchers?
LT// Can still employ the scientists in terms of analytic specialist services [KAB comment// services the route to RICHFIELDS having an income?]
CH// Public health is all going in the direction of placing the responsibility more on the population.
PieterF// How might a RI change the status quo, how can it help to analyse the data that is out there currently and being looked at by researchers already?

PaulF// Don't know what RQ to ask as unclear where the data is coming from. What dimensions of the data will be there irrespective if naturally occurring data or research-led data?
JL// Many products are flooding the market. Will there be a potential that all will be wearing a device and more of this data will be available in the future? Hence, these questions are important to talk through now.
LT// It could be that there become more differentiated groups regarding the amount of data they share or the concerns they might have over tracking and privacy.
MM// Might have to justify soon why one shouldn't use something such as fitbit rather than why should use it.
CH// Continuous data are useful. Is it possible to look at this data and map it to other things that have been going on (weather/political events etc.)? Also is it possible to follow the model of Michael J Fox Foundation and ask when someone purchases a fitbit if they would like to opt in or opt out of research. If so, then might allow the more structured capture of data.
LT// Opt in versus opt out is a discussion point e.g., UK NHS medical records have remained with the right to opt in.
[JdlC// Can we send literature to someone in RICHFIELDS for collation? It would be good to keep a record of all the papers that have been accessed by us all. Can participants all access Basecamp? KAB comment// Checked with Christina and can achieve this via Linked-In page].
RvdH// Paprika has the potentially to allow for cultural analysis with the same recipes in different languages, nationalities. It would be interesting to look at the international differences between the recipes. Similar app in the NL which also links recipes to one's shopping basket – linking what you cook at home and what you buy at the shop and vice versa.
CH// Paprika is an isolated app and not capable of providing that service. Could provide information on the recipes that are being searched with any medial or public health interventions.
LT// How much could be inferred from app searches?
GV// Food as eaten and have to translate into ingredients, perhaps looking at recipes can be useful to help with this preparation. Can add their own recipes.
PFrancois: can download on to multiple devices. activity of preparation and how recorded, on the go, or 2 nights a go etc.
MM// business data of website required for this.
CH// the search strategy to get the meaning of the searchers, where dwelled on certain recipes.
J, ES// activity on your computer, how deal with the machine.
CH// web based bhrial intervention conducted, even with pre-defined data, still hard to unpick what information had an effect.
PFrancois, eye tracking study. not big data.
AT// recipes structured in a standard fashion as make easier to search for ingredients etc.?
MM// imagine that some will be structured (list of ingredients) and some will be free text (procedure).
Government to retail.
J, NL: how link with the supermarket?
LT: businesses are obliged to provide information annually.
CH: Performance indicator measurement in terms of organic food.
PFrancios: simple notion of supply chain.

CH: economic growth of organic foods? belief organic food is inherently good? is this the point of this dataset. tenuous link to nutrition as all product based?
LT: what is the unit of analysis
IE: market share of products
PFrancois: DK marketing of organic foods? Volume. Don't know if also non-organic foods are recorded to calculate market share.
CH: more value available at lower cost, economic argument to push organic foods.
PFrancois: health metrics and economics metrics, link the two.
LT: observations trends issues.
CH: whatever outcome wish to look at.
JdlC: feedback:
Fitbit: social exclusion. define pop as only have IP. social background of users to interpret data. far from perfect crossing of data.
Can we use the data. Sell it. no one reads tandcs.
Food prep. household level.
Ethical issues.
CONDOR project.
J, NL: sales data see what you've purchased.
Q1U// What does this data allow you to do within your discipline? What research questions can it help answer (or allows you to ask) within your discipline?

Q2U// What unit of observation does it represent – who or what can you attribute this data to – the individual, household, an organization, geographical location, social interaction etc.?

- Aggregated
- Individual

Q3U// What are the shortcomings of the data?

- Validity
- Identifying population group

Q4U// How useful would this data be to a (discipline) studying food-related phenomena?

### *Group 2: Monique Raats (UK), reported by Siân Astley (BE)*

**Case 1: Ocado**

***1. What do these data allow you to do within your discipline? What research questions can the data help answer (or allows you to ask)?***

- Shopping behaviour, sales receipts.
- No public access to data, transcribe individually.
- Would have to have the cooperation of Ocado and or users to access data.
- These types of apps are no different to individuals donating their receipts.
- App owner, presumably, are undertaking analysis of the data for their business, e.g. promotions. What could RICHFIELDS offer the owner?

***2. What unit of observation does the data represent – who or what can you attribute these data to – individual, household, an organization, geographical location, social interaction etc.?***

***3. What are the shortcomings of these data?***

- Single or even multiple shopping events are not that informative without supporting meta-data since most households are not single occupancy.
- It is an assumption that food purchased is related to health.

***4. How useful would these data be to a (discipline) studying food-related phenomena?***

Might be useful to observe trends in purchasing of food types over time (e.g. plant-based foods). However, what do these trends imply?

There are two reasons researchers might want access to these data; (1) interested in seeing if there is something OR (2) know something and want to use the data to support this.
Would any conclusions drawn around these data be useful or better than existing methods where we know individuals lie or forget?

**Case 2: MyFitnessPal**

***1. What do these data allow you to do within your discipline? What research questions can the data help answer (or allows you to ask)?***

***2. What unit of observation does the data represent – who or what can you attribute these data to – individual, household, an organization, geographical location, social interaction etc.?***

- Geographical location
- Social interaction with groups

***3. What are the shortcomings of these data?***

- Users are motivated, which means they are not representative of the population as a whole.
- MyFitnessPal is limited in the same way as those completing paper diaries (ca. 4-10 days depending on motivation).
- Publically facing data are as unreliable as traditional methods (e.g. food diaries).

- In Copenhagen airport, wifi and Bluetooth track behaviours around cafes and restaurants. However, the ethics of using these data is debatable. Also, bias in behaviours around food is established and there are methods to deal with these.

***4. How useful would these data be to a (discipline) studying food-related phenomena?***

- RICHFIELDS's task is to connect data not dictate how it might be used.
- Breadth of data might be useful (e.g. a little from lots of people). Also, there would be information about how people are using these technologies.
- User would have to sign-up to share the data or have a deal with the App owner, but the role of RICHFIELDS is not to collect raw data. Nevertheless, to interrogate the data, raw data might have to be collected and processed (e.g. average weight can only be derived from weight).

To scenarios: (1) accessing data without the individual's consent (e.g. analysis of Twitter feed) and (2) directly engaging individuals or businesses to access data.

There are also Apps that merge food purchase data from different sources (e.g. MySupermarket) but only where the owners permit this (e.g. Ocado does not).

There is an issue of whether researchers want data from one source or many.

### *A3.3 Inference break-out session feedback*

**Breakout Session 2 (inference):**

- What do we need to know about data to make these useful for research of food related phenomena (e.g. intake, practices, health)?
- What do we need to be recording about the data in a meta-data document?

### *Group 1: Lada Timotijevic (UK), reported by Kerry A. Brown (UK)*

Fitbit case//
Trends, long term data, cross-reference with data to find certain cross-cutting dimensions, liquid community. No membership to enter and can leave.
Shortcomings:

- Children not included, inability to distinguish individuals or characteristics of pop as data attached to geographical IP address.able to use data? t and c's. ethics – privacy issues.
- Who produces/owns the data. Not everyone uses the application. individual, aggregator.

Paprika

- Cultural RQs. via language
- Public health interventions evaluation.
- Search strategy and computer recorded activity. how interact with app.

Shortcomings: data irrelevant in isolation.
G2R DK

- Supply chain.
- Assumptions? Organic = healthier? Economic? What is meant by organic
- What is being sold. What is being wasted?
- Stats bias likely?
- Context, market share, not individual level

Fitbit: e.g., cooking practices = anthropology; psychologist = mechanism of change; sociologist = social influence.
GV: representative of a population. quantitative data and structure so can do calculations. translate unstructured data into structured form. readable by machine. feedable into a code to allow stats. Also, validity (told truth) in some way, biomarkers to check diary data (also way biomarker collected/prepared).
AT: Calculate by BMR via than biochemical.
CH: how confident of data, +- 50% acceptable?
LT: what about having more than one fitbit
allergy app:if have enough in the survey then doesn't matter, wide enough data capture should iron out these?
PFrancois: size alone insufficient
JdlC: dating app populated by robots. need the data provenance.
LT// how can one validate?
JdlC// whole database, to see patterns than would not be possible by manual/natural development. Software has a poor measurement then have to go back and recalibrate all of the data.
GV: N=300 to calculate 95% p value.
CH// how externally reference though. Above only internal.
PF// not enough sensitivity to pick up energy balance change of 5%.

Pfrancois// aim of RICHFIELDS quality assurance? Provenance. Licensing. All failed, all passed, somewhere in the middle so that can help raise standards with new apps.
JdlC// no need to store data?
LT// standard setting body LT's aim of RICHFIELDS
CH// link data only if reach certain standards
AT// healthcare standards, minimum indication of data. stats market research panels, can criteria used by these panels inform some principles/standards
Paper idea: what are the pros and cons of being a quality assurance organization?
J DISH-RI// How has data been used before?
JdlC// Experts versus general public. experts decide what is required and what is quality, not only expert knowledge…age old discussion?
LT// how can data be interpreted by those from different perspectives?
PFrancois// scientific minimal standards, but can also accommodate different issues such as the monetary minimal standards.
LT// what counts as data will be different in different disciplines.
PFrancois// try to keep epistemology out of the project. broad consensus between scientists.
CH// health data has to be validated as responsibility to provide correct clinical diagnosis
DISH-RI// legislation of data source. IP legislation.
PFrancois// How can RICHFIELDS use existing licenses? Score on i) productivity levels (resources, time to clean, maintain data), don't want to over promise, how can potentially charge ii) levels of data quality – means different things to different people, knowledge of provenance, how well documented, change in databases, reporting structures, syntax validity, completeness of data, expert input, system incorporate new views iii) agility of system, upgrade in one software attached with new functionality how does this work with federated system, how adapt system. nested and broad level ontologies, potential schemes and new code languages, across 4 use cases, RICHFIELDS a 5th use case. high risk proposal and ability to fail spectacularly.
LT// structured or unstructured data?
PFrancois// both. one structured, and one other side that is unstructured and placing structure on it. DBPedia rogue scrape of Wikipedia (http://wiki.dbpedia.org/). Seshat: Global History Databank project in 2011 where imposing structure (https://evolution-institute.org/project/seshat/). Also lawyer/jury notes database imposing structure.
AT// CEN standards https://www.cen.eu/Pages/default.aspx
GV// Food provider/food navigator http://www.foodnavigator.com/Market-Trends/Big-data-project-set-to-reveal-consumer-food-habits-health
LT// define quality in phase 1?
MM// where comes from, how collected. no control over how people fill in the data. methodology valid. database driven apps, selecting food eaten from a pre-populated app, so need the database from these selections to be valid and then raise the validity of the app. reducing the systematic biases/errors in the app.
JdlC// indicators such as gender might not be pre-defined, gender changeable?
PF// make own food and no barcodes then cannot ensure uptodate databases.
LT// possible to link the data and infer motivation/why people do anything?

MM// data network created looks like could be useful. all on individual level, individualised interventions, decisions based on your own data. why some interventions fail. feedback on new achievements how does your network react to this data provision?
CH// need the history of change?
MM// can access this information, fitbit data is open and link to tweets and re-tweets, might not need to manipulate anything. don't see the sophisticated innovations in the research area as wait until everything is accepted as valid. barcode scanning used 2% as databases are considered insufficient in research but in weight loss commercially 60% use barcode scanning. how work in science vs to how work commercially.
CH// how know how data is changing in response to app developments? how track this.
MM// sensors not quite as good as could be in 5 years, who knows, still in commercial system as they are taking a chance on the evolution?
JdlC feedback
Q1// PF provenance of data.
Validity – objective. Sensors – calibration, tolerance?
Content issues
Minimal of standards – minimal criteria. Consumers wish correct data and have to say what is good scientific practice? 2 pages, different approaches for different disciplines. no knowledge meaning associated with data per se.
Purpose of the database required to link?

Q1// what do we need to know about the data to make it useful for research or food-related phenomena (e.g., intake, practices, health)?
Provenance
Validity – sensitivity, specificity, accuracy, reliability, degree can be generalised
Process of aggregation
Context
Continuity of data – kcal reductions due to data software changes, core data updates or data input changes?
Q2// what do we need to be recording about the data in a meta-data document?
James // approached myfitness pal not happy to be involved, wouldn't open up.
BK// Q quality crucial. who is going to give the certificate of quality, what body to check, who give the stamp, difficult?
A JdlC// RICHFIELDS frame of reference, RICHFIELDS give a stamp of quality?
LT// what is considered to be quality may differ across disciplines or interested parties. minimum of scientific standards representative of quality.
MMR// focus different across disciplines, nutrition data collected in psychology vice versa not of the same standard.
AT// declare or compulsory?
JdlC// co-ordinate the effort to obtain data enough to justify ownership of data? value is the co-ordination. country differences are okay.
PF// curating a service working with data providers.
BE// FOI asking for data in this way considered hostile in DK?

JdlC// was not considered hostile in ES
MMR// rather than 101 researchers asking for the same information from an organization, RICHFIELDS can allow access to this information via one point.

### *Group 2: Monique Raats (UK), reported by Siân Astley (BE)*

- Meta-data, the information placing the data of interest in context.
- Raw data, data for individuals, data for the population, etc.
- Need to know the question, e.g. what is available that might be manipulated to answer a given research question.
- Consumer related data – need to know about the individual
- Data sources, e.g. one app or many apps and aggregated data
- Backwards compatibility
- Monitoring over time
- Movement data using wearables
- Methods of calculation from raw data to output
- Need an API that is responsive to changes in Apps under analysis

***Where might we be in five years' time given the design must be future proofed?***
Is there any point building a RI if we have to get data from engaged individuals?
On the basis that most of the good app building is outside the research environment how does research capitalise on this?

Need to adapt sources that are available. For example, the UMCG cohort is being monitored intensively, but there are sources (e.g. electricity consumption) that might also be informative.

Good quality data are relative, e.g. nutritionist and psychologists have different needs. Quality and biases will differ amongst the sources, as do reasons for participation (e.g. can afford the latest gadget, requirement for insurance, external monitoring).

In order to create the design, do either the research question or the areas of interest need to be defined? Brainstorming with key research sectors to elicit the needs.

Food cannot be considered in isolation if the goal is health since other factors are also important. This is also important if RICHFEILDS is offering a reward for sharing data in the first place.

How in the meta-data is information about the individual being collected? Data capture meta-data is different from meta-data for individuals.

Choose a few Apps in a particular area and see what we can learn from the meta-data included in those sources. What is the range of data (un)available?

The research community is distant from the average user in terms of persuading them to share their data. Linking an individual between Apps is also an issue; shopping links to recipe use links to consumption of electricity/ gas, etc.

National differences in public health monitoring could reveal information. Similarly, national outlets might have data that would reveal more about consumer behaviours. Willingness to share might be cultural, e.g. Scandinavian countries might be more willing to share information for research.

Needs to be some sort of exchange to facilitate sharing. If we want to attract consumers, RICHFIELDS needs to explain what we are doing and how they will benefit, leading to an exchange of data. However, RICHFIELDS was intended to facilitate access not collect raw data. If the data are not being generated in the research field then researchers need to generate a case for access.

There might be a different response from sectors holding the data (e.g. government versus commercial sector), although industry increasingly is recognising that participating in research benefits their perceived role in society. Experience from the case studies in RICHFIELDS and other projects could help inform the process.

Defining the meta-data is too hard/ too wide/ not specific enough.

Where is funding for RICHFIELDS coming from as a functioning RI? And what are the needs of those funding it? Should we just try to do this with a few Apps? What do researchers want from RICHFIELDS?

Software engineer would appreciate standards and aggregation of data.

Researcher – meta-data and raw data, clean data, same ontology, same language, etc., which would support sharing of data.

How can the data in Apps be converted to something that might be useful for the research community rather than trying to infer something? i.e. provision of quality raw data although that might depend on who was using the outputs.

*3.4 Plenary discussion*

### *Breakout Session 1 (data), Group 1 reported by Javier de la Cueva*

⇒ What do these data allow you to do within your discipline? What research questions can they help answer (or allows you to ask) within your discipline?

FitBit:

- The app is related to the exercise you make and the diet
- How much do you move, comparisons between time
- Need long data to assess trends
- There is a way to cross data, far from perfect, but from this crossing you can arrive to certain criteria

Paprika:

- Language choice is huge – that means different cultures
- Doing/looking for recipes is not eating them
- Media or public health interventions, the gov. is giving messages so it would be in the mapping
- How you input the data
- Data, search strategies, recording the data in the computer

G2R Denmark:

- Supply chain
- Organic is healthier/economic? Made to make organic food cheaper
- Is it related to health?
- It is a trend issue
- Market share- can see market share of products
- It is only selling, not says what is wasted

⇒ What unit of observation does it represent – who or what can you attribute to these data to – the individual, household, an organization, geographical location, social interaction etc.?

FitBit:

- IP Address
- A kind of community (liquid)

⇒ What are the shortcomings of the data?

FitBit:

- Can we use the data?
- Can we access the data?
- If you look at the terms and conditions, they have the right to sell it – they have to care about the privacy

- Ethical issues: new social construct for science?; who owns the data, researchers for free?; privacy issues; opt in vs out – the right is to opt in and not opt out
- Trends are very useful but they are top level – there are trends in the intervention group but not in several groups
- Research infrastructure: is it simply reproducing he research issues in a high level?
- Difficulty: who uses the app; who produces the data
- Children not included
- It is not clear you can identify individuals – specific population, demographics; the higher the income the care of health
- Social exclusion – middle class tools, how to find the income of the users, they only have IP so they only have a place
- It is important to know the social background of the users of the application
- Forced consent/informed consent?

Paprika:
- Perhaps the context of shopping is important – obtaining the data in the purchase moment
- List of ingredients can be structured but not proceedings

G2R:
- Bias problems

⇒ How useful would these data be to a (discipline) studying food-related phenomena?

FitBit:
- They can link it to Twiter/Facebook and aggregate it
- Food purchase, depending on the type of family, food preparation changes
- Aggregate level, you can access the data individually, some people publish their data

G2R
- It makes sense in the context

### *Breakout Session 1 (data), Group 2 reported by Kristrún Gunnarsdóttir*

**Case 1: Ocado**
First point of call is accessing the data ...the business-model of it. How do you propose a model that has something in it for the grocer. Would it be possible that grocers are interested in a buy-in to an image of social-responsibility.

So who owns the data, is it the business that owns the data or the consumer, and how is that negotiated?

One idea is to have consumers press a button at the check-out to donate their purchase data. This perhaps relies on club-membership, like Tesco Clubcard, although, then you have to take into account that people purchase their food all over the place.

But moving away from the business model, what is the correlation between food purchases and health? Why would be want the data in the first place, what kind of information is inferred? Large datasets give you population-level information, how can this sort of data be acquired to understand individual food practices, family practices, singles practices and so on.

Concerning population-level studies, they can perhaps be very successful in getting large grocers on-board and individual consumers perhaps in large numbers, when the studies can be framed in the public good.

But for other types of studies, for example, sociological or behavioural studies, shopping habits can tell the part of a story but would the information value could be significantly improved by adding other forms of observation (digital data-driven or qualitative), tying together food purchases and health-related issues, food waste, not reporting truthfully, and many other aspects of a persons behaviour that are interlinked with purchase.

**Case 2: Calorie counter.**
Why would the researcher want this data? ...but we might need a system to bring the data in to contribute to a research agenda. We may not know yet where this sort of data fits in.

Limitations of the data, for example, it being collected while on a short-term diet. Are we interested in how long a person can hold it out to collect intense data on food consumption?

We need to access metadata and informational quality, create it from access to raw data. Richfield is not going to want to just collect and store a lot of data. On way of approaching such data, given one has access, use (time-line) visualizations to detect trends on which new research questions can be based.

Again, the issue came up about accessing data, say, why would MyFitnessPal want to share any of the raw population data they practically own. And, related to this is the difference between public

and 'private' share, painting oneself in one light or other, publicly as opposed to privately, but these biases are very well known in psychology and are showing up here in new guises. Say tweeting publicly about buying weightwatchers products, then tweeting to a friend that you ate 5 portions. Same with the link-up to Facebook.

There is also a question about tracking people's mac addresses in and around airports as an example of how to get away from the biases, as a more natural state observation, although, that doesn't quite help us in assessing how to make use of MyFitnessPal.

There is also a question of the extent to which Richfield as a research infrastructure provides data or is a service that can assist researchers in accessing and understanding data that would be useful to their research agendas… but the service is not yet defined since we don't know what services researchers might want.

### *Breakout Session 2 (inference), Group 1 reported by Javier de la Cueva*

⇒ What do we need to know about data to make these useful for research of food related phenomena (e.g. intake, practices, health)?

Fitbit:
Anthropologist – cooking practices

How representative is the data of the population

- Are these a select few respondents?

Validity of the data, objective measures – how can we trust the validity of the data?

- Assumes people are inputting true information – can we tease apart whether or not this is not a dummy data?
- Are the sensors calibrated? What is the tolerance? Or if you are comparing the data
- Someone could have 2 fitbits
- Brain mapping issue – tool to collect and calibrating issues
- Dating websites – fake profiles

Can set a minimum of standards – Richfields can offer a minimum approval; for 2 different samples we can offer 2 minimum criteria

- Consumers want correct data
- We have to state what is a good scientific practice
- Evaluation of minimum criteria, statistical market research – community is not the panel but minimum criteria
- Community is liquid, a panel not
- For what purpose my data set is used or what was used before, so to make links between research
- It can be two pages, different disciplines would have different approaches

- Quality managers, to validate
- Data has no knowledge meaning, so it is not in the traditional discussion experts – laymen
- The issue is go for this label
- Any standard has implications that come from the discipline
- There are some basis, of course, there is a broad consensus – anything can be data
- In the context of health you need measures
- The whole problem of obesity is an example of epistemology problem
- New metrics, new indicators. BMI as an indicator
- What is the legislation of the data? Potential stamp of quality and legal/ethical
- Levels of productivity, effort, data cleaning, maintaining. Levels of data quality means hugely different things – knowledge of provenance, changes on DB, legacy of old DB
- Layers of experts and their involvement
- We are speaking about behaviour and health values
- Data quality – how to control it. You cannot control what people insert
- Measure is very easy, sometimes the valuables are a choose-able set
- Individual fills as they understand, it is not selected from a commercial database
- Problems with categories – what is the motivation of people? Everything is now into individual level, it is not one suit fits all
- Fitbit is open so it can be linked to his – I don't need to manipulate anything
- Release early, release often
- 30 products Netherlands/9% products UK can be identified

⇒ What do we need to be recording about the data in a meta-data document?

- Clarify where data comes from
- Linked to calibration issues, we need to understand how it is obtained and organised
- Context of data – how much about the context of the "behaviour" is captured?
- Generalizability – how generalizable the data is? This requires understanding the sampling process.
- If studying aggregate data in order to identify trends - trends of what? It is a liquid community and the software is changeable – being able to know what is changing is crucial
- Show of turns – community and software

### *Breakout Session 2 (inference), Group 2 reported by Barbara Koroušić Seljak*

- Research questions need to be defined
- Then we can specify an 'ideal' app to define meta data
- We could make a pilot in which we would collect API specifications for the selected apps from areas defined in Phase 1 – to semi-automatically define meta data
- We need to know what are our plans for 2018 – who is out main stakeholder?
- Paulo: business – health – biomarkers, endpoints
- Paul/Maud: clean, aggregated data
- Meta data are needed for data quality validation
- Meta data needs to support adaptive-ness
- Complex meta data – how it is collected by the app
- Focus groups to define meta data
- What information to provide as feedback to consumers/users
- Ontology is a must

## Annex 4: Workshop evaluation

An online feedback form was sent to all participants but only nine responded, even after several reminders (45%).

| Response | Chart | Percentage | Count |
|---|---|---|---|
| Research | | 66.7% | 6 |
| Industry | | 11.1% | 1 |
| Government | | 0.0% | 0 |
| Consultancy | | 22.2% | 2 |
| Other | | 0.0% | 0 |
| | | **Total Responses** | **9** |

### 1.a. In which disciplines are you an expert?

| Response | Chart | Percentage | Count |
|---|---|---|---|
| Biological and Medical Sciences | | 14.3% | 1 |
| Chemistry and Material Sciences | | 0.0% | 0 |
| Earth and Environmental Sciences | | 0.0% | 0 |
| Engineering and Energy | | 0.0% | 0 |
| Humanities and Arts | | 0.0% | 0 |
| Information Science and Technology | | 14.3% | 1 |
| Physics, Astronomy, Astrophysics and Mathematics | | 14.3% | 1 |
| Social Sciences | | 28.6% | 2 |
| Other, please specify... | | 28.6% | 2 |
| | | **Total Responses** | **7** |

### 1.b. In which disciplines are you an expert? (Other, please specify...)

| # | Response |
|---|---|
| 1. | Dietary assessment in population study |
| 2. | exposure risk analyses |

### 1. Date of this event

| Response | Chart | Percentage | Count |
|---|---|---|---|
| Very satisfied | | 77.8% | 7 |
| Satisfied | | 11.1% | 1 |
| Neutral | | 0.0% | 0 |
| Unsatisfied | | 11.1% | 1 |
| | | **Total Responses** | **9** |

### 2. Venue (Exchange Avenue, Schiphol, NL)

| Response | Chart | Percentage | Count |
|---|---|---|---|
| Very satisfied | | 88.9% | 8 |
| Satisfied | | 0.0% | 0 |
| Neutral | | 11.1% | 1 |
| Very satisfied | | 0.0% | 0 |
| | | **Total Responses** | **9** |

### 3. Meeting room

| Response | Chart | Percentage | Count |
|---|---|---|---|
| Very satisfied | | 88.9% | 8 |
| Satisfied | | 0.0% | 0 |
| Neutral | | 11.1% | 1 |
| Unsatisfied | | 0.0% | 0 |
| | | **Total Responses** | **9** |

### 4. Lunch

| Response | Chart | Percentage | Count |
|---|---|---|---|
| Very satisfied | | 33.3% | 3 |
| Satisfied | | 55.6% | 5 |
| Neutral | | 11.1% | 1 |
| Unsatisfied | | 0.0% | 0 |
| | | **Total Responses** | **9** |

### 5. Duration

| Response | Chart | Percentage | Count |
|---|---|---|---|
| Very satisfied | | 66.7% | 6 |
| Satisfied | | 33.3% | 3 |
| Neutral | | 0.0% | 0 |
| Unsatisfied | | 0.0% | 0 |
| | | **Total Responses** | **9** |

### 6. More generally

| Response | Chart | Percentage | Count |
|---|---|---|---|
| Very satisfied | | 66.7% | 6 |
| Satisfied | | 33.3% | 3 |
| Neutral | | 0.0% | 0 |
| Unsatisfied | | 0.0% | 0 |
| | | **Total Responses** | **9** |

### Comments

| test comments |
|---|
| Thank you for the opportunity to participate in a very interesting research topic |
| The workshop was well organized and managed. I enjoyed some interesting conversations with colleagues from a range of disciplines. |
| It would be fantastic to have regular meetings to be updated on how the project is going on |

### 1. Applicability of topics

| Response | Chart | Percentage | Count |
|---|---|---|---|
| Very satisfied | | 66.7% | 6 |
| Satisfied | | 33.3% | 3 |
| Neutral | | 0.0% | 0 |
| Unsatisfied | | 0.0% | 0 |
| | | **Total Responses** | **9** |

### 2. Lecturers

| Response | Chart | Percentage | Count |
|---|---|---|---|
| Very satisfied | | 33.3% | 3 |
| Satisfied | | 55.6% | 5 |
| Neutral | | 11.1% | 1 |
| Unsatisfied | | 0.0% | 0 |
| | | **Total Responses** | **9** |

### 3. Interest in topics

| Response | Chart | Percentage | Count |
|---|---|---|---|
| Very satisfied | | 88.9% | 8 |
| Satisfied | | 0.0% | 0 |
| Neutral | | 11.1% | 1 |
| Unsatisfied | | 0.0% | 0 |
| | | **Total Responses** | **9** |

### 4. Depth of coverage

| Response | Chart | Percentage | Count |
|---|---|---|---|
| Very satisfied | | 55.6% | 5 |
| Satisfied | | 44.4% | 4 |
| Neutral | | 0.0% | 0 |
| Unsatisfied | | 0.0% | 0 |
| | | **Total Responses** | **9** |

### 5. General

| Response | Chart | Percentage | Count |
|---|---|---|---|
| Very satisfied | | 66.7% | 6 |
| Satisfied | | 22.2% | 2 |
| Neutral | | 11.1% | 1 |
| Unsatisfied | | 0.0% | 0 |
| | | **Total Responses** | **9** |

### 6. BREAKOUT Session 1: Measurement

| Response | Chart | Percentage | Count |
|---|---|---|---|
| Very satisfied | | 44.4% | 4 |
| Satisfied | | 44.4% | 4 |
| Neutral | | 11.1% | 1 |
| Unsatisfied | | 0.0% | 0 |
| | | **Total Responses** | **9** |

### 7. BREAKOUT Session 2: Inference

| Response | Chart | Percentage | Count |
|---|---|---|---|
| Very satisfied | | 44.4% | 4 |
| Satisfied | | 44.4% | 4 |
| Neutral | | 11.1% | 1 |
| Unsatisfied | | 0.0% | 0 |
| | | **Total Responses** | **9** |

### Comments:

| Response |
|---|
| Well conducted group and comprehensive report in topics more and more important |
| I thought both the breakout sessions produced a range of feedback directly relevant to the aim of RICHFIELDS. |

### How did you first find out about this workshop?

| Response |
|---|
| I was contacted by the organizer |
| Invitation |
| Through WP4 colleagues |
| I was invited from the project's management |
| I am part of the RICHFIELDS project team |
| I'm a project partner |

### How could we have improved this meeting?

| Response |
|---|
| Just arranging new meetings |
| Bit less psychology |
| It would have been nice to see a greater variety of disciplines represented |
| The next meetings should probably be focused on a specific topic. Every topic covered is so wide that it is enough to cover an entire meeting |
| Nothing springs to mind |
| By involving more stakeholders |

### What was the most interesting about this meeting?

| Response |
|---|
| The knowledge of tools to help citizens in managing diet-physical activity-health topics is crucial to understand if and then gaps occur and following the evolution to exploit the tools at the best level possible |
| Seeing what is going on in the field and share ideas |
| Feedback from Stakeholders. |
| Meeting people from other professional areas |
| Hearing the different perspectives across the differing domains |
| Brain storming with stakeholders. |

### What was the worst about this meeting?

| Response |
|---|
| Nothing |
| Nothing |
| I think the workshop could have been improved by inviting a greater number of stakeholders from a wider range of disciplines. |
| Nothing |
| Nothing springs to mind |
| Nothing to be complained about. |

### Do you wish to continue receiving news and updates about RICHFIELDS?

| Response | Chart | Percentage | Count |
|---|---|---|---|
| Yes | | 88.9% | 8 |
| No | | 11.1% | 1 |
| | | **Total Responses** | **9** |

### Would you be interested in attending future RICHFIELDS stakeholder workshops?

| Response | Chart | Percentage | Count |
|---|---|---|---|
| Yes | | 88.9% | 8 |
| No | | 11.1% | 1 |
| | | **Total Responses** | **9** |