CRŒSS

CROSS-domain & ESS (European Statistical System)

Topic title: Innovative tools and sources for diary-based surveys
– Household Budget Survey and Time Use Survey
– Test and Maintain

B5460-2020-INNOVTOOLS-HBS-TUS

Authors

Statbel – Social statistics – Kelly Sabbe

Destatis – Jerome Olsen, Eric Bitzer, Elke Nagel, Birgit Lenuweit, Daniel Knapp, Christiane März, Anja Sommer, Pablo Brecht

Hbits – Joeri Minnen, Pieter Beyens, Enak Cortebeeck

Contact information

Project coordinator – Statistics Belgium

Kelly Sabbe

Statbel, the Belgian statistical office

Koning Albert II-laan 16

1000 Brussels

+32 (0)2 277 66 30

kelly.sabbe@economie.fgov.be

Beneficiary – Destatis (Statistisches Bundesamt) (DE)

Jerome Olsen, statistician Gustav-Stresemann-Ring 11 65189 Wiesbaden Jerome.olsen@destatis.de	Eric Bitzer, statistician Gustav-Stresemann-Ring 11 65189 Wiesbaden +49 (0)611 75 8482 eric.bitzer@destatis.de	Birgit Lenuweit, senior statistician Gustav-Stresemann-Ring 11 65189 Wiesbaden +49 (0)611 75 8572 birgit.lenuweit@Destatis.de

Subcontractor – MOTUS developer

Joeri Minnen

hbits CV – Spin Off Vrije Universiteit Brussel

Witte Patersstraat 4

1040 Etterbeek

Tel: +32 497 18 95 03

Joeri.Minnen@hbits.io / Joeri.Minnen@vub.be

Reader's guide through the document

In this report, we often refer towards our prior collaboration and report, SOURCE™ (Software Outreach and Redefinition to Collect E-data [Through MOTUS]) (Minnen, 2020).

Overview of the Work Packages and deliverables

Start with the **main document** to get a **comprehensive overview** of the entire project or go the **Work Package 1 (WP1) – Deliverable 1.2 (D1.2) - Administrative final executive summary** for the highlights.

If you want to know more about the **definition of a cross-domain data collection platform** for the European Statistical System (ESS), take a look at WP2. D2.1 gives an overview of the **necessary research components** to establish this cross-domain data collection platform, while D2.2 gives an overview of the **technical documentation of the database and API**.

WP3 describes the development of new Time Use Survey (TUS) & Household Budget Survey (HBS) components, so called **microservices**. These microservices consist of the **geolocation microservice** (D3.2) and the **receipts scanning microservice prototype** (D3.3). To make the developed microservices accessible, **R-cran packages** were developed for distribution (D3.4).

Next the aforementioned developments and building on the feedback we received during our SOURCETM project from the pilot test within Statbel and Destatis and within 28 other National Statistical Institutions (NSIs), **new interfaces for TUS and HBS** were defined (WP4). The **user interface (UI) and user experience (UX)** of MOTUS went through extensive feedback rounds (D4.1). Destatis and StatBel held **pilot tests** with the new interface, the findings can be found in D4.2. New **user manuals** were developed, to make the new interface more comprehensive (D4.3).

WP5 gives an overview of the **roll-out of a data collection platform** (MOTUS) within the NSI, which has been done by Destatis. Some important aspects were documented, such as the **database technology pillars** (D5.1), the **installation report of the MOTUS containers** (D5.2) and a report on the **vulnerability tests** Destatis conducted (D5.3). Finally, the existing Common Statistical Production Architecture (**CSPA**) documentation was updated.

Conducting a pilot test with MOTUS

The default language of MOTUS was English for the test study but the default language can be defined at the beginning of the construction of a research. So if you only have a study in German it is no problem at all. Before a **pilot test** can take place **in a country's own language**, **translation support** is needed for the documents mentioned in D2.1 – research components and D4.1 – UI/UX for TUS and HBS. To conduct your own pilot test, please contact Joeri Minnen.

Once all the documents are translated, you have the necessary **user manuals** (D4.3) translated, which consists of a user guide **for the respondents** (front-office) and a user guide **for the NSI/researcher** setting up the survey (back-office), where all the possibilities of MOTUS are explained.

CONTENTS

List of Figures

List of tables

WP1 – Coordination and communication

Objectives

The first work package dealt with the coordination, financial management and communication during the project. WP1 also included the contact and follow-up (in tasks and financial) of the subcontractor since the project proposal involves MOTUS.

---

1   WP1 – D1.1 – Administrative interim executive summary

Overall goal: description of the intermediary results, progress

---

The interim executive summery was submitted halfway through the project. The administrative final executive summery gives the final overview.

2   WP1 – D1.2 – Administrative final executive summary

Overall goal: description of the intermediary results, progress and SWOT.

2.1   Project Context

The main responsibility of Eurostat is to provide reliable and comparable statistical information to the institutions of the European Union (EU). An important instrument to achieve this is through the promotion of – as far as possible – harmonised statistical methods across its member states.

In this view the European Statistical System (ESS) has been built up gradually. While Member States (MS) are responsible for the collection of the data and the compiling of the statistics for National and EU knowledge building, it is the role of Eurostat to support the ESS to grow out as a network that strives to search to harmonized procedures. Two important data collection methods within the ESS are the Time Use Survey (TUS) and the Household Budget Survey (HBS). Today they are collected on the basis of a gentlemen's agreement between the participating country and Eurostat.

TUS is used to get a grip on the daily activities that people do and how these activities are imbedded into the contextual display of time, place and social environment. TUS is harmonised via the HETUS-guidelines (Eurostat, 2020). Today 3 rounds of data collections have been organized in Belgium, the 4th round started May 1st 2022. A third version of the guidelines have been published in 2019. In 2030 TUS will enter the IESS agreement on an optional level (The European parliament and the council of the European Union, 2019).

HBS collects in a large detail households expenditures on goods and services. In this way, the survey gives a picture of the living conditions in the EU. There is a strong continuous effort towards more comparability and output harmonisation in HBS, but there are still important differences in terms of frequency, timing, content or structure. HBS is performed by each Member State to calculate weighted macroeconomic indicators used for national accounts and consumer price indices. Eurostat publishes output since 1988 and this with intervals of 5 years (Eurostat, 2022). The last waves are from 2010 2015 and 20. In 2026 HBS will enter the IESS agreement (The European parliament and the council of the European Union, 2019).

2.2 Need for analysis

The outcomes of both the TUS and HBS survey are being used for many different purposes and used in several contexts. For TUS the databases are the basis for gender statistics, working times, media use, to calculate the collective rhythms and cultural differences between countries and to calculate trends over time to name only a few of the possibilities. For HBS the primary use is the calculation of expenditure habits of a population in a year's time, but also other topics such as the question of social protection, poverty or even domain-specific insights related to transport, energy, education, health and many more are studied based upon these datasets.

A wide range of (international) institutions make use of both the TUS and HBS data, from policy orientated organisations to NGO's, to universities and even increasingly private organisations.

Most importantly is that these data have the capacity to monitor our rapidly changing society. The need to monitor has been recognized by the 2011 DGINS Wiesbaden Memorandum (DGINS, 2011). The Memorandum however also stressed that better data is needed, both in terms of coverage and comparability.

Most surveys are suffering from a lower respondent participation. In most countries, every member of the household is asked to keep a 2-day diary for TUS and a household jointly keeps a record of their consumption during 14 days for HBS. Also, a household and individual questionnaire is asked to be completed.

Another initiative of Eurostat is the discussion with the Task Force (TF) TUS and HBS on the Governance Guide. This initiative aims to define an operational model how countries can make use of a tool, but also can ask for modifications or the development of new components. More and more the discussion upon the establishment of a platform comes to the surface. Especially this last requirement is an important step towards a more efficient data collection process for MS.

To counter the decreasing response rates and to keep on providing essential data, Eurostat is supporting the modernisation of TUS and HBS via technological developments. These developments should lead to:

1. an improved responsiveness of the respondent,
2. a better integration of other sources of information, and
3. a more efficient data collection process for MS.

Many initiatives have been taken already by Eurostat to give an answer to this request, and also in relation to TUS and HBS. The underlying grant is the 4th generation but there is also the ESS-net project where MOTUS (Modular Online Time Use Survey) as a TUS tool is evaluated by the participating countries. To do so, a number of test studies are designed and carried out. Some new developments achieved in the CRŒSS project became available within the time frame of the ESS-net, the ESS-net participant were invited to test new versions of MOTUS and microservices. The same is true for other projects like the 'Innovative Data Collection Solutions for HBS and TUS'-project where VUB-hbits developed a geolocation Proof-of-Concept (PoC) as subcontractor of Sogeti.

2.3 Relevance of the project

TUS and HBS are two data collections in social statistics that are funded under the European Statistical Programme. Since 2016 consecutive grants focused on the transition to an online data collection strategy. In a second order also other data sources were mapped on this modernisation vision.

The grant round of 2018 projects included the SOURCE™-project (Minnen, 2020). In the SOURCE™ grant, the goal was to introduce, to document and to test MOTUS. Requirements as

1. functionality & maintainability,
2. tool reusability,
3. online availability,
4. usability,
5. user friendliness,
6. accessibility, and
7. data comparability

were used to evaluate MOTUS and to define critical developments.

The SOURCE™ report lists up the actions to be taken to include the HBS data collection in MOTUS, as well as the critical developments to upgrade MOTUS to an ESS-platform. The main element in favour is the powerful back-office of MOTUS that consists of different modular research components that support the Build, Collect and Process phases of the GSBPM-architecture. As a logical continuation of the SOURCE™ evaluation, the new consortium proposed the CRŒSS project.

2.4 Concept and methodology

"A shared knowledge is a strong knowledge".

This saying has been the slogan during the SOURCE™-project, and it was the same in the CRŒSS project, but with fortified insights.

Statistics Belgium, Destatis and hbits have a longstanding expertise in the collection of TUS and HBS diary data. During the SOURCE™-project a lot of information was documented and the software architectures to collect TUS and HBS of MOTUS, Statistics Belgium and Destatis were evaluated. Also, other NSIs were asked for input during TF/WG and bilateral meetings. At the end of the project a demonstration of MOTUS was given and evaluative input from 28 NSIs was asked.

Good critics were received on the modularity of the platform and the criteria of reusability and shareability. This is mainly due to the back-office of MOTUS and the underlying builders (e.g. device builder, survey builder, communication builder, research builder, language builder, invitation builder, dashboard builder, data builder and quality builder).

The consortium, and the other NSIs were also asked to give negative remarks and to look to future possibilities for the data collection platform. These inputs led to 3 task descriptions:

- Action list to upgrade MOTUS to HBS.
- List of critical requirements for TUS and HBS.
- Evaluation of the UI/UX front-office of MOTUS.

These task descriptions fuelled the strong question from the consortium and the NSIs to upgrade MOTUS to a data collection platform that can handle TUS and HBS and that is easy to be shared. Both objectives are included in the CRŒSS project.

2.5 European added value & impact on non-EU countries

The entire setup of the CRŒSS project was focused on:

1. the need for better information in terms of coverage and comparability;
2. the call for better support to the respondents;
3. the call for better support to the NSIs.

The CRŒSS project introduces a cross-domain data collection platform for TUS and HBS that is easy to share. At the same time, it will introduce new ways of data collection. This introduction itself will be innovative through the inclusion of microservices. The future of social statistics is insightful and flexible on the one side, but also rigorous and comparable on the other side. The MOTUS data collection platform is able to merge these requirements in one environment.

MOTUS will become easy to share, cooperation between countries will be supported and the need for comparability will be boosted. Information will be more structured and the data availability will be immediate. The benefits of the CRŒSS project will be present for all European Countries, as well as for the EU-bodies.

The impact of the project will also support the cooperation between Europe and other continents and international institutions. Already today Eurostat receives large attention for the developed Inventory of Tools and Sources (Eurostat, 2021), and representatives of Eurostat are invited to present their work and views on meetings of the UNSD, UNECE and others.

2.6 Framework of the objectives

The MS experiences and expertise were collected via a questionnaire in 2018 and an Inventory was made to share this knowledge (Eurostat, 2021). The Inventory shows which components are serviced by the various tools:

- Does the tool have a web app, a mobile app, a back-office, an API, …?
- How does the tool apply to the criteria of maintainability, reusability, online availability, usability, user friendliness & accessibility, and data comparability?

The Inventory is available separately for TUS and HBS tools.

Currently 15 TUS tools and 10 HBS tools are documented. The focus now is to improve the reusability of the tools, to enlarge the support of the collection of data and to improve the harmonisation strategies. These elements are present in the MOTUS platform, having a front-end, a back-office and a server environment. All components are split up according to the 3-tier principle for software architectures.

2.7 MOTUS – Modular Online Time Use Survey

MOTUS is a software platform developed by the Research Group TOR of the Vrije Universiteit Brussel (Belgium). Its primary aim is the online collection of diary based information. At its origin lies the collection of Time Use Survey (TUS) data. Time diary research tries to capture the actual behaviour (What people do) of people within its context (Why people do it).

MOTUS combines a back-office and a front-office. The back-office supports the researcher to collect and disseminate data. The front-office is available to the respondent to take part in the studies.

The back-office of MOTUS incorporates four phases: design, collect, analyse, advice.

Figure 1: Main GSBPM-phases in which MOTUS provides support
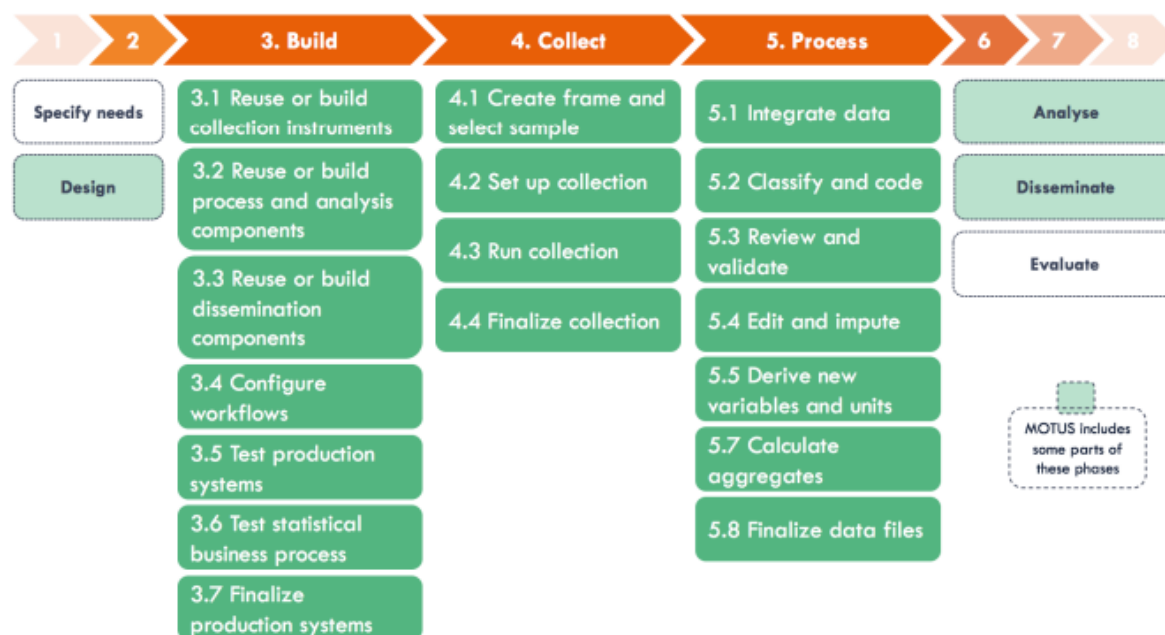
Each phase includes a number of builders available to the researcher to run a data collection. The use of builders supports MOTUS in its most powerful asset: modularity. It is the composition of the builders, and the choices being made within these builders that define the actual set up of a particular research. As such, MOTUS makes it possible to define multiple researches, that can run at the same time, even within the same respondent (for panel research purposes).

The respondent can make use of a web application that runs in a broad range of web browsers and a mobile application available for Android and iOS. All information is shared and synchronized. Information on respondents can also be attained from other data sources and from sensors that are connected with MOTUS. To support the data collection and the interaction between devices specific server capacities are implemented.

A first data collection took place in 2013. Since then MOTUS was further expanded, and more studies were executed through MOTUS. In 2019 the first data collection using passive data input via sensors was realized.

Due to the functional and technical setup of MOTUS other diary based surveys, like the HBS can also highly benefit from the business processes available in MOTUS.

2.8 Objectives of the CRŒSS project

The overall objective of the CRŒSS project was to create a cross-domain platform for TUS and HBS that can be shared in the ESS and with a positive impact for respondents, researchers and NSIs. The two general objectives are divided over more specific objectives.

The more general goals of the CRŒSS project were:

1. to define MOTUS as a cross-domain tool for both TUS and HBS data collection,
2. supported by a platform that is easy to share and which inherently supports data comparability on the ESS-level,
3. an improved User Interface (UI) and User Experience (UX) of the MOTUS applications,
4. the inclusion of external data via microservices (e.g. geolocation ...),
5. the positioning of the respondent in the middle of the data collection strategy and
6. to cope with the privacy and security regulations, now and in the future.

The CRŒSS project followed upon the SOURCE™-project (Minnen, 2020). The previous project's goal was to get to know MOTUS as a platform for the data collection of TUS data. This project was realised by the consortium Statbel and Destatis, and with the input of in total 28 different National Statistical Institutes (NSIs). The subcontractor was VUB-hbits, a research group with expertise in TUS, data collection and the developer of the MOTUS-software platform.

The SOURCE™-project evaluated the capacities of the MOTUS-software platform and came to the conclusion that MOTUS can grow out to a cross domain data collection platform. In this respect the CRŒSS project (as a composition of CROSS and ESS) had 2 innovative goals:

1. to upgrade MOTUS from a tool for TUS to a tool that also includes HBS, and
2. to define a prototype of an ESS-platform by using the container technology.

By means of the CRŒSS project the consortium wanted to address the full intentions of the 2011 Wiesbaden Memorandum to support respondents, researchers and NSIs in their efforts to provide Trusted and Insightful Statistical Data. In second order this project gives a high value to the criteria of reusability and shareability (DGINS, 2011).

The project had 5 Work Packages (WP) which had clear mutually reinforcing and innovative goals.

Table 1: Overview of the work packages

WP number	WP Title	Lead beneficiary
WP1	Coordination and communication	Statistics Belgium
WP2	Definition of a cross-domain data collection platform	Statistics Belgium
WP3	Development of new TUS & HBS components (microservices)	Destatis
WP4	Defining and testing of new interfaces for TUS and HBS	Statistics Belgium
WP5	Roll-out research data collection platform	Destatis

Figure 2: Work Package overview
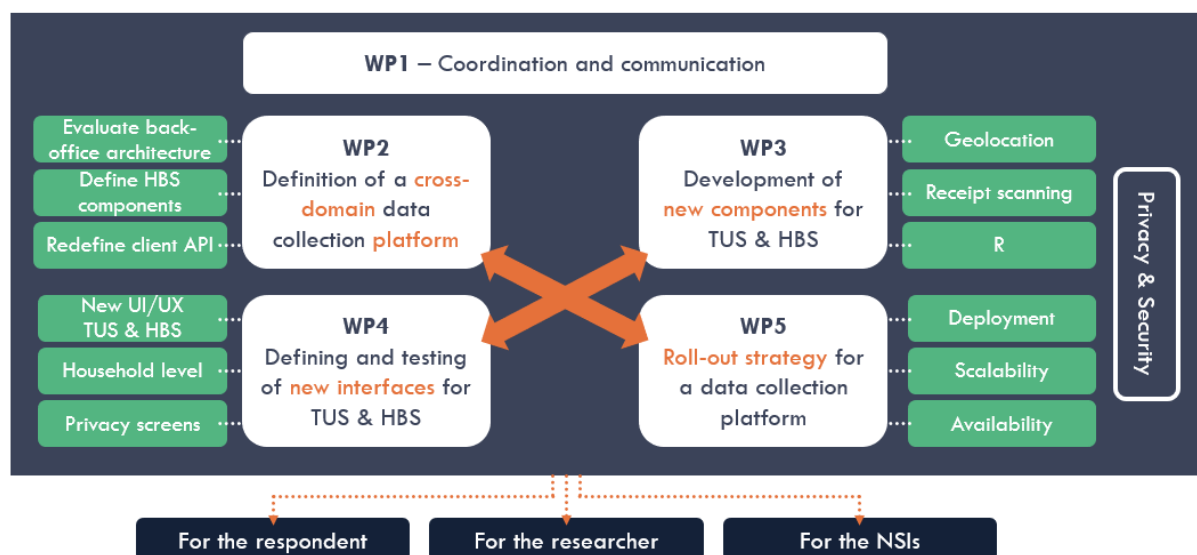
The designers of the project deliberately have chosen a small consortium, but with a large exposure to the other NSIs. This means that the outcomes will be communicated after the peer-review of the results that will take place after October 31st and the documentation is detailed, but also that during the project the interaction with the other NSIs was seen as a key element. Therefore, all information, documentation, presentations and code gathered and created during this project will be made public.

2.9 Overview of the main achievements

The CRŒSS project included a number of different activities in order to arrive to a prototype of an ESS platform that is able to collect TUS and HBS data. This work departs from the capacities of the MOTUS data collection platform.

Central to the back-office are the "MOTUS-builders". The use of these builders supports MOTUS in its most powerful asset: modularity.

Figure 3: MOTUS builders for TUS and HBS according to GSBPM
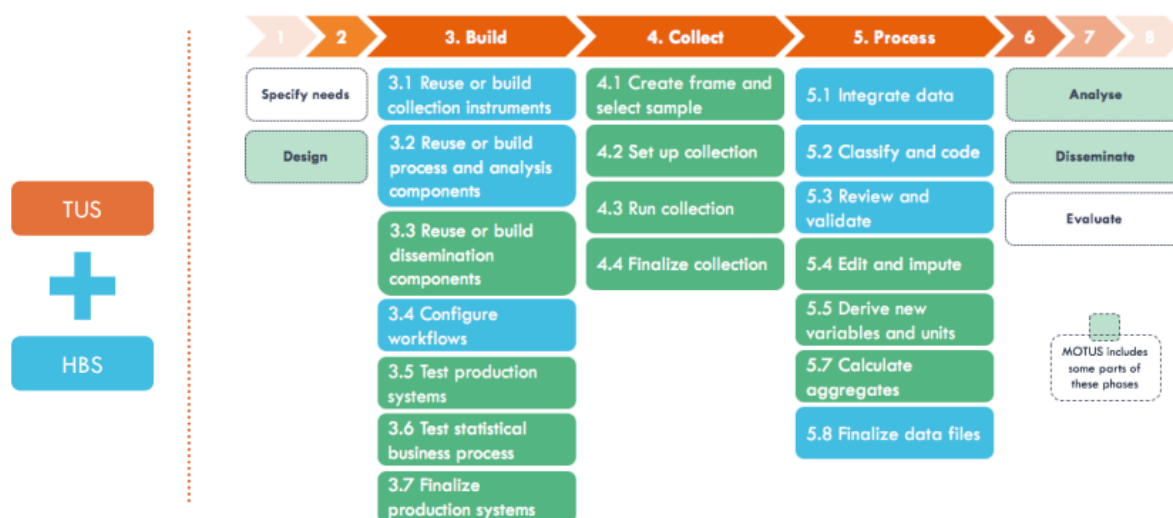
The front-office is available to the respondent to take part in the studies, as it holds a mobile application for smart devices running on Android and iOS, and a web application to be used in a browser. All information is shared and synchronised between the applications.

The architecture of the MOTUS-software platform is realized as a 3-tier software architecture:

1. having a client tier (user interface),
2. an application tier (functional process logic) and
3. a data tier (data storage).

Figure 4: 3-tier software architecture of the MOTUS platform
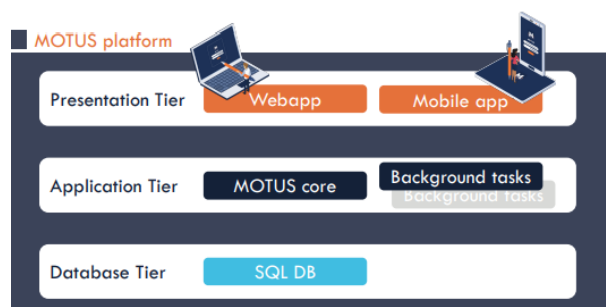
The SOURCE™-report documented the critical developments for the platform to achieve an ESS-standard, and showed the development story to include HBS to MOTUS, reusing already existing parts of the TUS methodology. This gained knowledge provides the basics for the new grant (Minnen, 2020). By means of the CRŒSS project the consortium wanted to improve and upscale the possibilities of the MOTUS-platform, in terms of reusability and shareability.

2.9.1 WP2 – Definition of a cross-domain data collection platform

State of play

A cross-domain tool is beneficial for researchers and NSIs in an organisational perspective. To date the Inventory shows no tool that services both TUS and HBS (Eurostat, 2021).

Added value

The objective of WP2 was to upgrade the MOTUS-tool from a tool for TUS, to a tool that can service both TUS and HBS.

Actions

One of the main characteristics of MOTUS is its modularity, by means of the different research builders. These different builders defined for TUS (12 in total) can be reused for HBS. This makes the step-up to incorporate HBS into MOTUS more than feasible. To upgrade MOTUS to an HBS-platform the main task was to define a new diary builder for HBS.

Objectives

The specific objectives were linked to the enlarging of the back-office of MOTUS to include HBS, besides TUS. This project reuses the MOTUS components or so-called 'builders' that are already available. To include HBS the definition of a new diary builder (specifically to HBS) was important. The technical components were already available at the start of the project.

WP2 focused on the definition of the different underlying parameters of the new diary component and the reuse of the other components or builders to define a HBS-research.

The (iterative) testing of the input ran in accordance to WP4 'Defining and testing of new interfaces for TUS and HBS'. The specific objectives of WP2 were:

- Definition/evaluation of the back-office architecture of MOTUS to include HBS
- Definition of HBS research components and research flow in MOTUS
- Redefinition of the client API

Conclusion: upgrading the MOTUS-platform by including a HBS diary

Towards other statistical domains TUS and HBS have the use of a diary in common. On the other hand the diary for TUS and HBS is different. By including a new diary component for HBS, and by reusing the already existing components of MOTUS, the upgrade of MOTUS to a cross-domain platform has succeeded.

Special attention was given to the database structure and the inclusion of the COICOP list in combination with an efficient search algorithm.

Figure 5: Database structure of the HBS diary
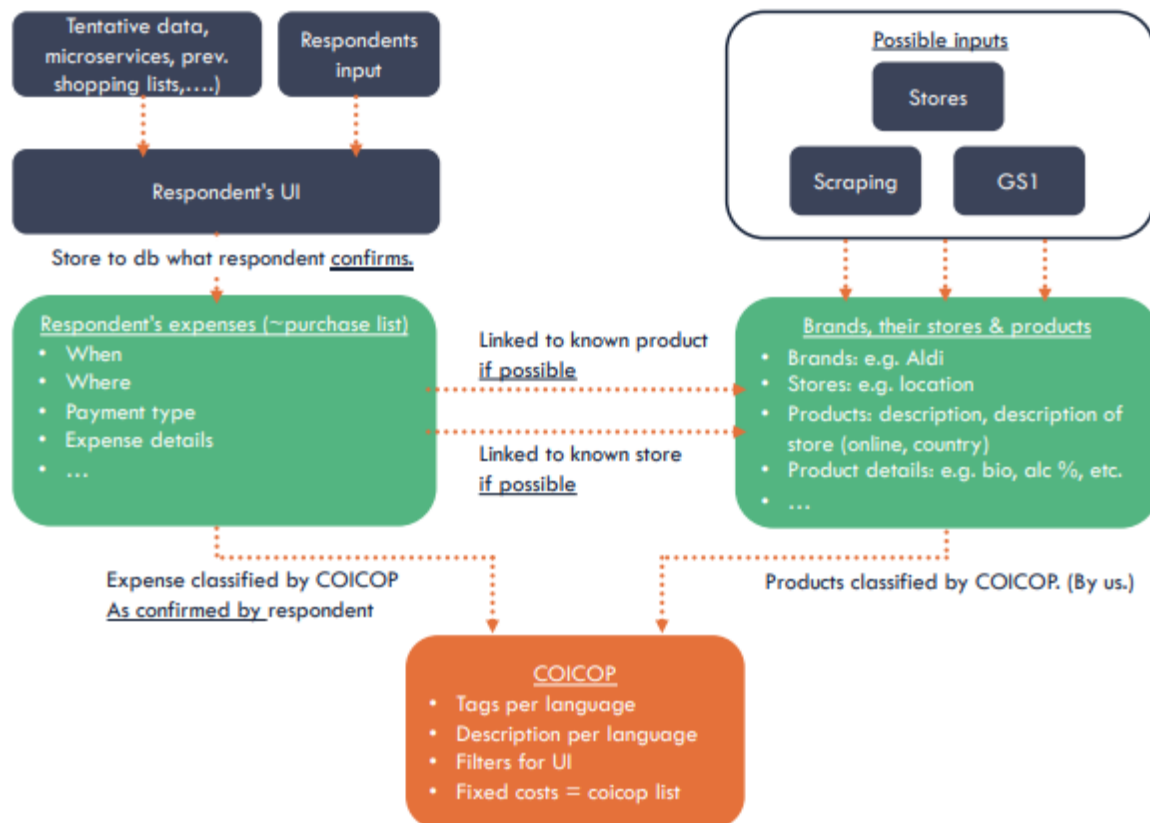
2.9.2   WP3 – Development of new TUS & HBS components (microservices)

State of play

Smart data can provide more detailed/accurate information on activities and purchases. For TUS and HBS some prototypes became available but needed to be evaluated and did not (yet) have the status of a microservice.

Added value

By including microservices, respondents, researchers and NSIs become better supported. Microservices can gather smart data in favour of TUS and HBS. These microservices enlarge the capacity and functionality of MOTUS while they are also equally shareable to other platforms and purposes.

Actions

This project developed a software architecture incorporating microservices that is linked to the MOTUS data collection platform, but which is equally reusable and shareable to other platforms by means of the APIs. The project includes:

- geolocation,
- receipt scanning and
- R-cran packages

as microservices.

The connection with microservices is a long-term strategy. Different microservices can be attached to support the TUS and HBS data collection for respondents, researchers and NSIs. It is also a profound way to connect wearables or other smart devices in the future.

Objectives

In WP3 three new functionalities (microservices) are included. Two of the new functionalities integrate smart data to enrich the primary data and/or support the respondent in the data collection process. Through smart devices information can be collected which is otherwise more difficult to include.

1. The first is the geolocation for TUS.

Figure 6: MOTUS microservices
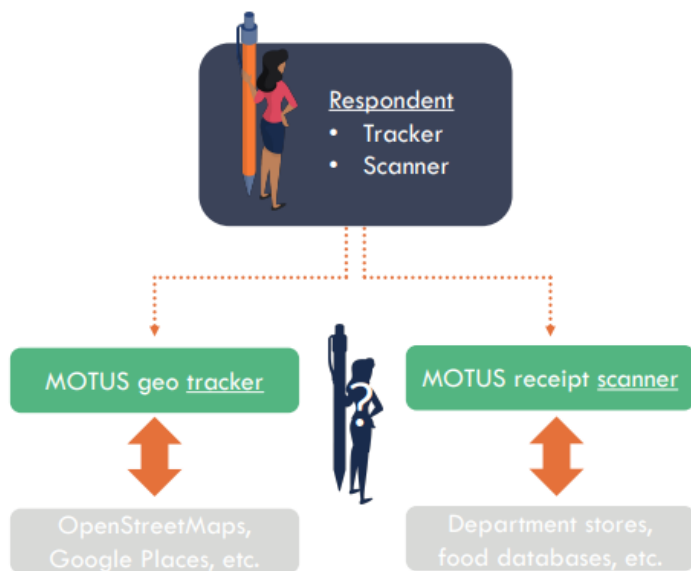
2. The second is the receipt scanning for HBS.

Figure 7: Receipt scanning for HBS
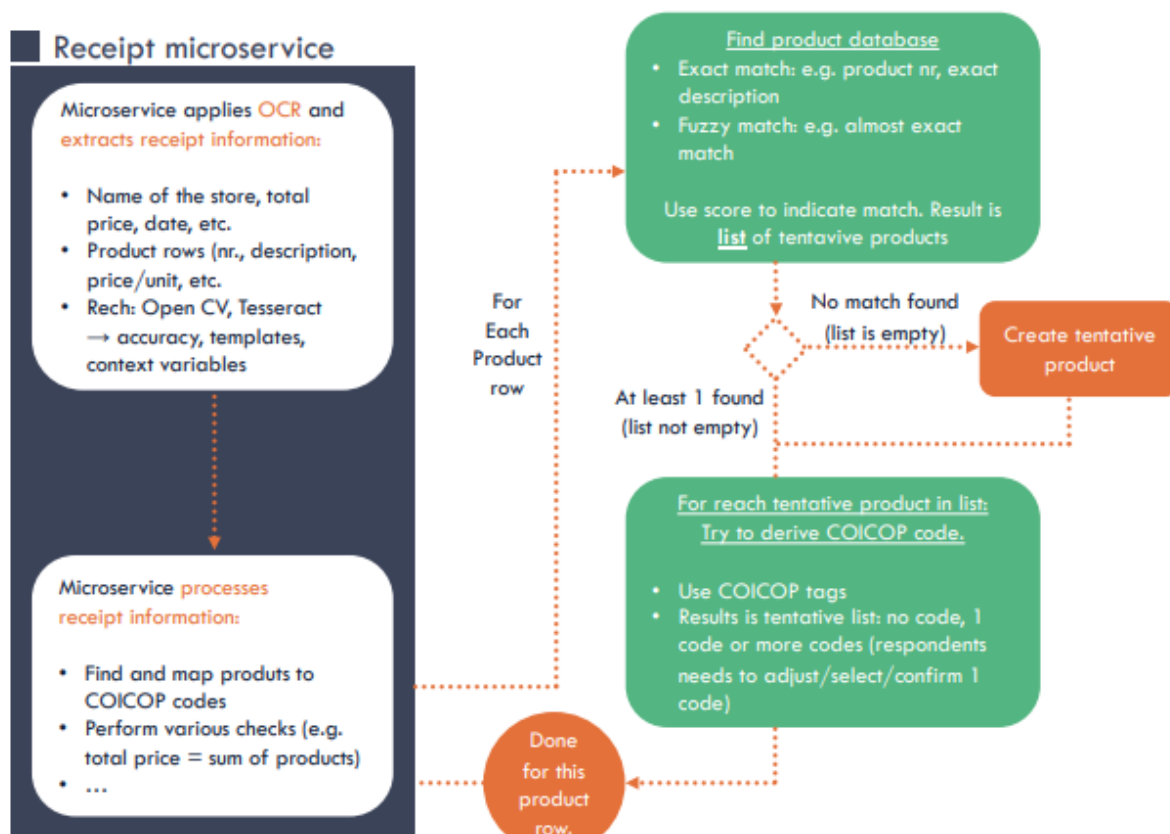
3. A third functionality (microservice) is the development of an R-cran package for TUS and HBS.

Figure 8: R cran package for TUS
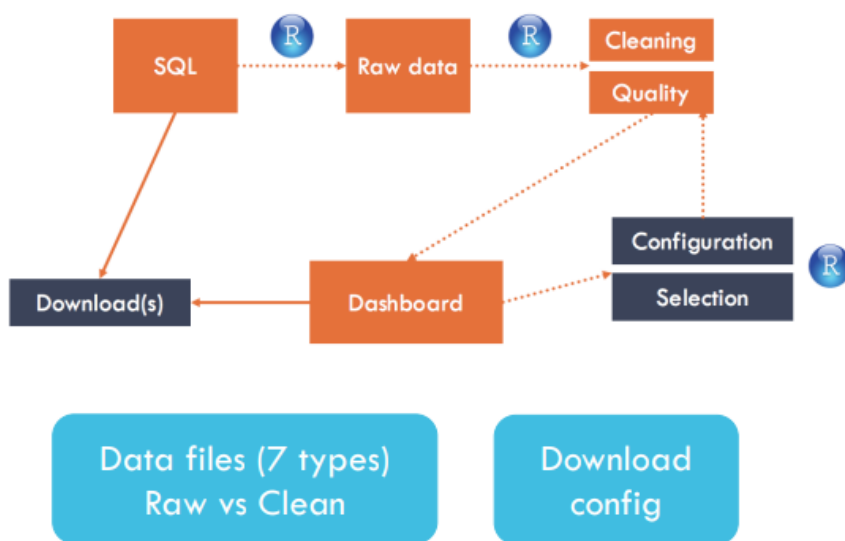
Through R-scripts various implementations can be developed to support the data collection and cleaning (quality checks, control) of the data. To include these microservices a shareable software architecture is needed, also comprising a database and API.

The specific objectives of WP3 were:

- Definition and development of a software architecture for microservices.
- Development and integration of microservices to support the data collection, for TUS and HBS.

Conclusion: including smart data to MOTUS by means of microservices

Microservices can handle various sort of plug-ins in a standardised way, including the geolocation (for TUS) and the receipt scanning (for HBS) plugin. By using microservices, smart data or passive data can be included in a data collection strategy. The microservices run independently, having an own database and API to support reusability and shareability.

Figure 9: MOTUS platform shareability
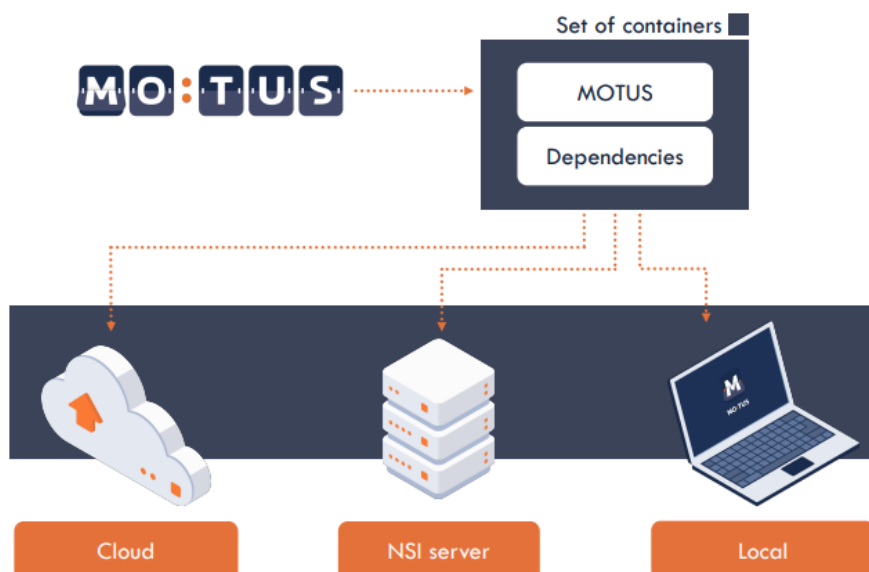
2.9.3 WP4 – Defining and testing of new interfaces for TUS and HBS

State of play

MOTUS was tested during the SOURCE™-project. Before the CRŒSS project no application covered both TUS and HBS, nor was able to let respondents interact with smart data.

Added value

By means of a 1-to-n strategy and the underlying client API, MOTUS is preconfigured to service both TUS and HBS surveys.

Actions

The main task was to evaluate new designed interfaces that connect to TUS and HBS studies and that are able to interact with data captured through microservices. Special attention was given to privacy-preserving screens and consent asking screens. The new design also needed to take into account the household as surveying entity.

Objectives

The TUS front-office (interface) has been tested in the SOURCE™-project (Minnen, 2020). Remarks on the usability of the TUS application were documented. The good aspects need to be preserved, the negative remarks to be solved.

WP4 covered the existing issues while it also included new innovations and developments that have an impact on the UI/UX of the front-office. The goal was to support the respondents in their data collection. A new interface is designed prior to this project. In this project the interfaces were tested and evaluated. Modifications can still be made afterwards.

The specific objectives of WP4 were:

- Update/new UI and UX of the MOTUS app for TUS and HBS.
- Include household level.
- Include privacy-preserving screens.
- Basic usability tests of the various elements (internal test, focus groups (with internal NSI colleagues, …).

Conclusion: usability testing

The UI and UX of MOTUS was already evaluated in the SOURCE™-project. Positive and negative remarks were collected and documented. In this project the usability of MOTUS for TUS were re-evaluated based on a modified UI and UX, and a new HBS UI and UX was tested.

2.9.4   WP5 – Roll-out research data collection platform

State of play

Eurostat as well as other international bodies encourage software developers/owners to document their tools according to the CSPA-catalogue (Eurostat, 2021). The CSPA has 3 levels:

1. a conceptual level,
2. a logical level,
3. an implementation level.

Extra attention is given to the criteria of reusability, shareability and comparability. MOTUS has been already documented on the conceptual and logical level.

Added value

Through the modular setup and the availability of both a front- and back-office, MOTUS is seen as a shareable tool. During the CRŒSS project this shareability was further extended:

- by including HBS (WP2),
- by including new functionalities (WP3) and,
- by including a new UI and UX design (WP4).

By including a roll-out strategy in WP5 the aim was to improve the shareability of MOTUS within the ESS, taking into account today's security and privacy requirements.

Actions

The project had the objective to develop a prototype where the installation of the MOTUS-platform, implemented through containers. This way MOTUS is easy to share and can be integrated into the NSIs own production line. The MOTUS-code runs on the 'application tier' of an NSI by means of the Docker technology. This way the development team of MOTUS can further develop/update and maintain the platform while at the same time comparability and shareability are maximized.

In essence the container technology provides four pillars that are essential for a reusable, shareable, flexible and maintainable ESS platform. These pillars are:

1. Deployment, meaning: remote/cloud deployment, fast installation, increased error reduction and an organised validation control, …
2. Scalability, meaning: multiple databases, replication, many/large projects in size and numbers of respondents at the same time without loss of performance, strong improvement of both analysis and transaction functions to accept a large number of short online transactions, …
3. High-Availability, meaning: database replication and replication of front-end connectivity to deal with system failures, …
4. Privacy & security, meaning: strict separation of databases, role management, …

Objectives

The second overall goal was to make the MOTUS data collection platform for TUS and HBS easier shareable amongst NSIs in order to support them in their data collection strategy. In the SOURCE™-project 4 different roll-out strategies were defined and evaluated (Minnen, 2020). The conclusion was to further investigate the container technology. This way an interpretation is given to the Governance Guide that Eurostat is defining together with the joint TF TUS and HBS.

The specific objectives of WP5 were:

- Define an efficient roll-out strategy to share MOTUS.
- Define the requirements to host MOTUS and integrate it to an existing production line.

Conclusion: documentation for the respondents, researchers and NSIs

The outcome of the project is related to the respondents, researchers and NSIs. The central position of the respondent is the combining factor, having an impact on all 3 levels. Test documentation, test plans and guidelines were developed for NSIs and researchers in how to setup MOTUS and how to work with it.

The main achievement is the concretisation of a data collection platform that can handle TUS and HBS data collections, which is easy to share and which by means of the roll-out strategy leads to comparable output data.

A comprehensive overview about MOTUS, who is behind it and how it works can be found online: https://www.MOTUSresearch.io/

Figure 10: MOTUSresearch.io
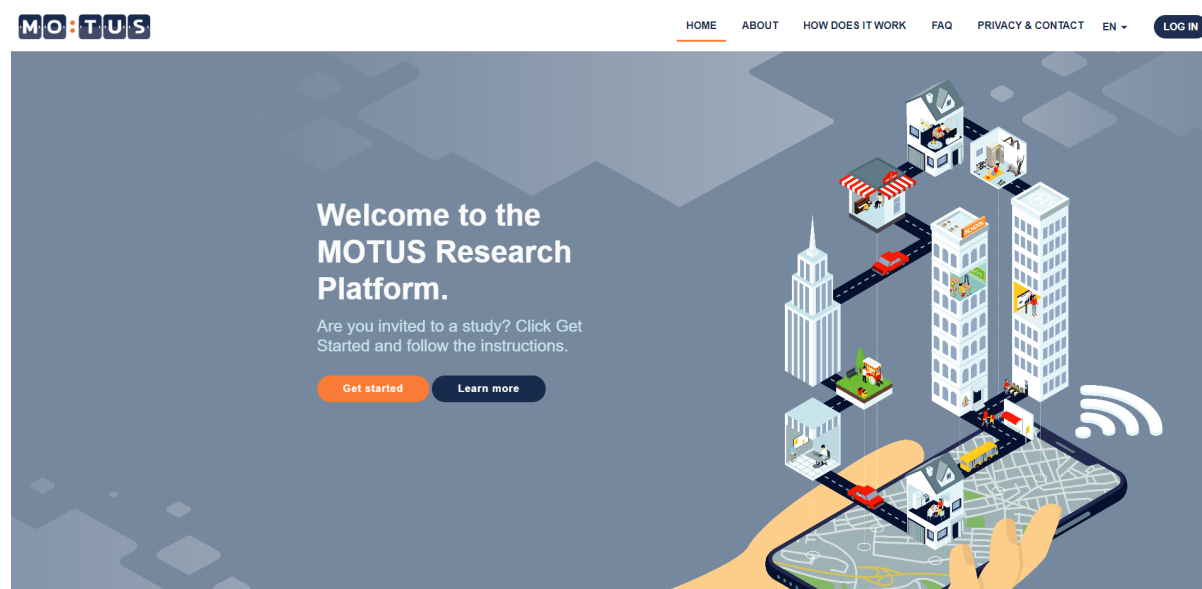
The website does not only provide user manuals, but also the FAQs, information about privacy and contact information. The same information can be found in store and in app. All information is available in English (default language), Dutch and French.

2.10 Impact of the project

MOTUS has achieved the ambition to define a cross-domain statistical platform (for TUS and HBS) that is easier to be shared between NSIs taking into account current requirements on privacy and security. The impact of the innovative solutions should lead to an improved situation for respondents, for researchers and for NSIs.

Many countries are planning a TUS and HBS data collection in the coming years. Time is short and at the same time the development of a tool/platform is a high risk for NSIs. A data collection platform that is able to organise automatised fieldworks, that makes use of common knowledge and comparative components and which still give the possibility to variate in country specific requirements is not only a short term necessity, but also a necessity on the longer term.

The long-term vision of MOTUS is fortified because the platform can host more statistical domains and innovative data collection strategies can be inclined in a reusable and shareable manner.

The CRŒSS project and the underlying MOTUS-platform is future proof since each of the WPs focuses on the three roles that were identified in the 2011 DGINS Wiesbaden Memorandum (DGINS, 2011):

- Respondents will get a new UI/UX and will be supported via smart data input, will have control over their data and will be able to learn from their own data. It will lower the registration burden of data.
- Researchers will be supported in building a research, in the data collection, and the preparation of the data, all from one software platform for TUS and HBS. It will enforce the comparability of the data.
- NSIs will have one supported platform available for multiple research domains which improves the management tasks. It will support the planning (in time and budget) and dissemination tasks of NSIs. It will collaboration between NSIs better possible.

The software architecture of MOTUS is focussed on a long-term usage of the platform, since the platform:

- Puts the respondent in the central position of the data collection process.
- Is developed according the 3-tier software architecture.
- Will be able to be hosted by NSIs through the container technology.
- Supports – through its technical setup – cooperative developments (e.g. NSIs, hbits).
- Is suitable to include today's privacy and security requirements.
- Facilitates a future without the – necessary – use of interviewers (contactless research where every step is automated but via a structured personal communication).

Figure 11: MOTUS architecture - UUID communication
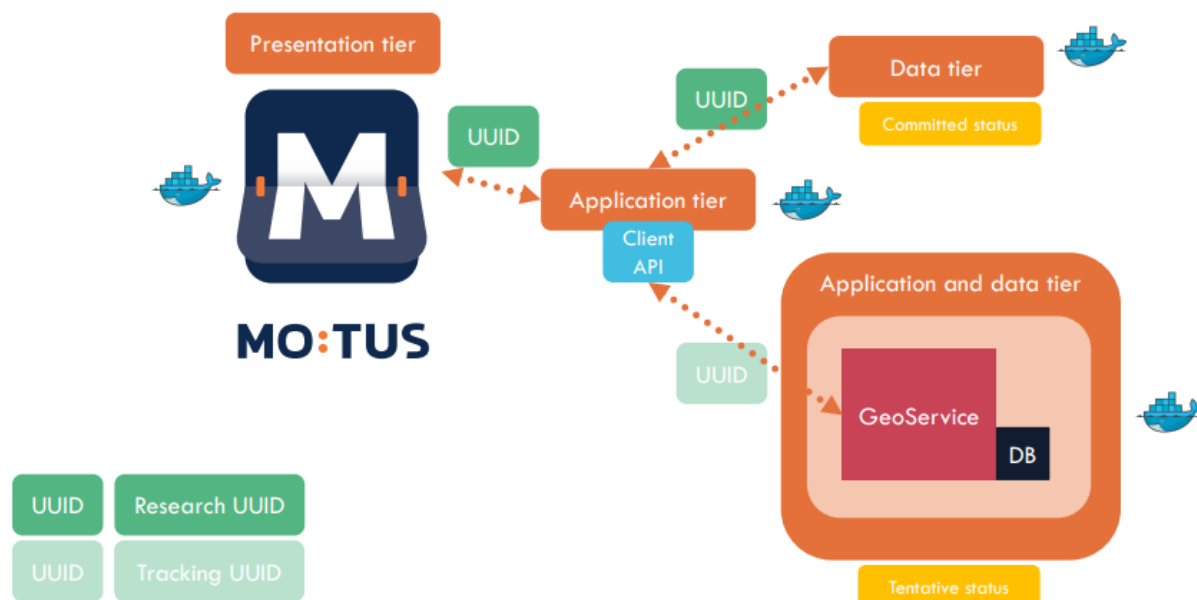
2.11 SWOT analysis of MOTUS: a summarized feedback

On the next page a summarized feedback is given, shown as a SWOT-analysis.

A SWOT-analysis gives an interesting view on the capacity of MOTUS to become a platform to collect official statistics in Europe and on an international level. We first started with a SWOT analysis during the SOURCE™ project, and based on the end conclusions the CRŒSS project was defined.

The goal of the project was MOTUS to develop/redefine to a Cross domain platform available to the European Statistical System. The cross-domain criteria was met by adding HBS to the already TUS-related platform. Both methods also make use of questionnaires. The ESS level is met by creating new deployment strategies to share MOTUS, the most flexible solution is the virtual deployment via the container technology. The new project tackled many of the listed opportunities and weaknesses at the end of the SOURCE™ project. Which now leads to a new SWOT overview.

Also of importance is that the (remaining) list of opportunities and weaknesses still meets the development strategy of MOTUS and so will become part of MOTUS in the near future. Some weaknesses related to the methodology of TUS and HBS itself but are crucial for MOTUS and TUS/HBS in general to become/stay/remain an accepted ESS-method. To do so, (as an example) TUS and HBS needs to become less burdensome for the respondent, and less costly for NSIs.

The weaknesses are important shortcomings of MOTUS today. Some points need software development, others belong more in the atmosphere of communication. These points need urgent attention. The threats have less to do with the platform as such, but are an essential part in order to get MOTUS accepted as an ESS-platform. And so of importance. MOTUS has by design a privacy and security requirement included. Nevertheless it is up to the national and international countries and levels to make their evaluation and, subsequently, to the development team of MOTUS to respond to this in a positive way.

Threats are to be seen as outside decisions that hamper the functionalities within MOTUS, and so that MOTUS becomes less efficient. An example is the GDPR-legislation but also security aspects that are essential to become part of an international architecture. These threats need to be monitored and to be solved as much as possible.

Strengths
External origin
• Cross-domain platform
• Laravel based platform
• Laravel based components (microservices)
• Cross-design apps
• Cross-design back-office
• Skin and brand definition
• Questionnaire functionality
• Diary functionality
• Household grid functionality
• Automatic communication
• Automatic task orientation
• Automated research flow
• Automated fieldwork follow-up
• Multi-device / Multi-platform
• Multi-language
• CMS-system – back-office
• CRM-system – back-office
• Modular components
• Modular research flow
• Standardization
• Reusability
• Comparability
• Shareability
• Scalability
• Service deployment
• Native deployment
• Virtual deployment via container strategy
• 1-to-N strategy
• Dynamic application build
• No third party components (like Google)
• APIs (input and output)
• CSPA-documentation
• GSBPM-documentation
• TUS-eco system (survey + diary)
• HBS-eco system (survey + diary + COICOP + Fixed costs)
• Connection to microservices
• API development
• Committed & Tentative mode
• Immediate data availability
• Privacy & security by design
• 3-tier architecture
• User role definition
• Penetration tests performed
• Accessibility tests performed
• Informative website
• Guideline to respondents
• Export compatibility
• Libraries

	Weaknesses
External origin	- Not yet a training facility for researchers - Not yet an informative website - Not yet a well-balanced guideline for researchers - No yet an online help-desk - No yet an immediate feedback - Impact of multi-mode effect - Not yet an interviewer role - Not yet integration of paper-and-pencil mode strategy - Growing gap between groups of individuals - Steep learning curve - Burdensome registration - Length of the participation
	Opportunities
	- Lower cost & time investment - Further improved built-in quality assessment - Country specific adaptations - Improved role management - Acceptance of other sources of information - Acceptance of administrative data sources - Respondent roles (e.g. personal interest) - No or less need for interviewer capacity - Combination with an interviewer - Improved and broader inclusion of R + TUS R-packages - Better compatibility with SPSS, SAS, … - On the go research - International interest - Multi-disciplinary - Other cross-domain opportunity (HBS, transport, tourism, …) - Sustainable Development Goals - Employability during pandemics or energy crisis - Community building, panel opportunity
	Threats
	- Growth of configuration requirements - License strategy not yet defined - Governance model not yet defined - System integration/harmonization - National laws - External privacy issues - External security issues - External ethical issues - Changes in the underlying software platform (Ionic, React, Flutter,…) - Old, not updated devices - Old, not updated browsers - Stability and size of development team - Server capacity - Stress test not yet executed

3   WP1 – D1.3 – Workshop

Overall goal: demonstration of the work done in the CRŒSS project. Report on the workshop and presentations.

3.1   Introduction

Originally, there was a workshop planned that would solely focus on the CRŒSS project. However, due to the fact that the previous workshop on innovative tools and sources that was planned by Eurostat in 2020 couldn't take place, we were asked if we would agree upon a collaboration with Eurostat. This collaboration meant that the workshop was linked to a joint Task Force TUS/HBS and that Eurostat helped with the organisation. Within the timeframe of the workshop the 4 projects of this innovative tools and sources round were presented, as you could see in the agenda.

The workshop took place on the 20th and 21st of September 2022 in Statistisches Bundesamt, Gustav-Stresemann-Ring 11, 65189 Wiesbaden. The workshop was hybrid, with the grant coordinators and Task Force representatives present in the room and all other interested participants (such as Working Group representatives) could follow online. In total, there were about 30 people present in the room and 50 participants online.

Workshop info, agenda and PowerPoint presentations are available on the Eurostat wiki page: https://webgate.ec.europa.eu/fpfis/wikis/display/ISTLCS/Workshop2022.

The presentation of the CRŒSS project is also available on our Webpage ITNRW: "NRW connect extern" as B5460_BE_1.3_Workshop_Presentation_ENG.pdf

- The workshop presentation contains the slides presented by Statbel, hbits and Destatis, as well as a live demo about the new UI/UX.

During the course of the project, we have updated the leaflet with general information about the project that was made by Eurostat during the SOURCE™ project. The leaflet of the CRŒSS project is also available on our Webpage ITNRW: "NRW connect extern" as B5460_BE_1.3_Workshop_LEAFLET_ENG.pdf. We distributed this leaflet during the Task Force on innovative tools and sources for HBS/TUS the day after the workshop.

3.2   Reflections on the presentation of the CRŒSS project

The presentation, demonstration and discussion of the CRŒSS project (2020-BE-INNOV) took place the first day at 15:15. We had to stop the presentation at 18:00 for the social event. We needed to include a beak and there were also a lot of interesting questions. This meant that there was not enough time to give the entire presentation (the demonstration of the R-cran package – D3.4 R-cran package microservice) or to give the participants time to try the app out, for which the opportunity was foreseen.

3.3 Wiesbaden Workshop Agenda

Wiesbaden workshop on Innovative Tools and Sources for Household Budget Survey and Time Use Survey: Agenda

Table 3: Agenda of the Wiesbaden workshop on innovative tools and sources for HBS and TUS

	Tuesday 20 September | Session 1 | 13:00-18:00 **Chair: DESTATIS**	
	Opening DESTATIS will welcome the participants and inform them about the content, objectives and practical arrangements of the workshop. Speaker: Lucas QUENSEL-VON KALBEN (workshop organiser, in the room).	**13:00**
1	**Introduction** Eurostat will set the stage providing an overview on progress done so far in the innovative tools and sources project. Speaker: Anne CLEMENCEAU (workshop organiser, in the room).	**13:15**
2	**Presentation, demonstration, discussion of Project 1 (2020-HU-INNOV)** The grant project team will present the main results of the project focusing on the project's outcomes. Speakers: Gergely JÓZSEF (in the room) and Tünde HAGYMÁSY (coordinator, in videoconference).	**13:45**
	Coffee break 30 min	**14:45**
3	**Presentation, demonstration, discussion of Project 2 (2020-BE-INNOV)** The grant project team will present the main results of the project focusing on the project's outcomes. Speakers: Kelly SABBE (coordinator, in the room) and Joeri MINNEN (team member, in the room).	**15:15**
	Coffee break – 15 min!	**16:45**
	Presentation, demonstration, discussion of Project 2 (2020-BE-INNOV) The grant project team will present the main results of the project focusing on the project's outcomes. Speakers: Kelly SABBE (coordinator, in the room) and Joeri MINNEN (team member, in the room).	**16:45**
	End of session 1	**18:00**

	Social tour and dinner	**18:30**
	More info about the social dinner is available on the wiki page.	

	Wednesday 21 September | Session 2 | 09:00-12:00	
	Chair: DESTATIS	
4	**Presentation, demonstration, discussion of Project 3 (2020-NL-INNOV)**	**09:00**
	The grant project team will present the main results of the project focusing on the project's outcomes.	
	Speaker: Barry SCHOUTEN (coordinator, in the room).	
	Coffee break – 30 min!	**10:30**
	Presentation, demonstration, discussion of Project 3 (2020-NL-INNOV)	**11:00**
	The grant project team will present the main results of the project focusing on the project's outcomes.	
	Speaker: Barry SCHOUTEN (coordinator, in the room).	
	End of session 2	**12:00**

	Lunch break – 1 hour	**12:00**

	Wednesday 21 September | Session 3 | 13:00-17:00	
	Chair: DESTATIS	
5	**Presentation, demonstration, discussion of Project 4 (2020-BG-INNOV)**	**13:00**
	The grant project coordinator will present the main results of the project focusing on the project's outcomes.	
	Speaker: Georgi MARKOV (in videoconference).	
6	**Call for proposals on Trusted Smart Statistics**	**14:00**
	For information	
	Eurostat will inform about the new call for proposals on "Trusted Smart Statistics: methodological developments based on new data source" (SMP-ESS-2022-TSS-METH-TOOLS-IBA) that was published in July 2022.	
	Speaker: Konstantinos GIANNAKOURIS (in the room).	
	Coffee break – 30 min	**14:30**
7	**Panel discussion**	**15:00**
	Panelist will discuss the results of the four projects and their wider impact, as well as provide indications about the way forward. Panelists will be invited to tackle concepts such as shareability, comparability, privacy, governance, trust.	
	Speakers: Anne CLEMENCEAU (Eurostat), Lucas QUENSEL-VON KALBEN (DESTATIS), Patrick LUSYNE (STATBEL), Florian KEUSCH (University of Mannheim), Bella STRUMINSKAYA (University of Utrecht).	
	Panel moderator: Eniel Ninka (Eurostat)	
	End of the session 3	
	Concluding remarks by DESTATIS and Eurostat	**16:50**
	Closing of the workshop	

4 WP1 – D1.4 – Technical report

> Overall goal: overview of the action plan and evaluation. Plan for the communication of the results.

4.1 Overview of the action plan and evaluation

The table below was constructed during the SOURCE™ project, to give an overview of what was available in MOTUS and an action plan of what could be developed in the future, based upon the SWOT-analysis. The developments during the CRŒSS project are in blue.

Table 4: To do's for MOTUS towards the future

#todo	Description	Available
Device builder e-HBS		
#1	Database structure (for e-HBS and e-TUS)	✓
#2	New web app build with HBS functionality	✓
#3	New mobile app (iOS & Android) build with HBS functionality	✓
#4	API connection HBS external plugin(s)/microservice(s) to client API	✓
Questionnaire builder e-HBS		
#--	Questionnaires before/after diary	✓
#--	Extra questions in relation to the consumption in the diary	✓
Diary builder e-HBS		
#5	Define COICOP-classification structure	✓
#--	Define extra questions in diary	✓
#--	Define parameters in diary	✓
Questionnaire builder e-HBS		
#--	Define on screen information	✓
#--	Define emails	✓
#--	Define notifications	✓
Device builder e-HBS		
#6	Include geolocation/geofence plugin - POS sensing	
#7	Include iCARD plugin – POS purchases	
#8	Include receipt scan plugin - scan and OCR analysis	
#9	Include PSD2 plugin – personal bank account	
Language builder e-HBS		
#--	Multilingual System buttons	✓
#--	Multilingual Content researches	✓
#--	Multilingual Communication settings	✓
#--	Default language database	✓
#--	Default language meta-data	✓
Research builder e-HBS		
#--	Creation of a stage	✓
#--	Assignment of a task	✓
#--	Assignment of communication	✓
#--	Assignment of triggers	✓

#--	Assignment of actions	✓
Invitation builder e-HBS		
#--	.csv/.xlsx file functionality	✓
#10	External API to register as participant	✓
#--	UUID creation	✓
#--	Username creation	✓
#--	Password creation	✓
Dashboard builder e-HBS		
#--	Follow-up fieldwork	✓
#--	Follow-up respondents/households	✓
Data builder e-HBS		
#--	Download datasets as .csv, .sav., .RData	✓
Quality builder e-HBS		
#11	Extension MOTUS-Package for R to clean/prepare the data	✓
#12	Definition of quality criteria	✓
Compute builder e-HBS		
#13	Extension MOTUS-Package for R to analyze the data	
Visualization builder e-HBS		
#14	Extension MOTUS-Package for R to visualize the data	
User Interface e-HBS		
#--	Technical functionalities web and mobile app	✓
#15	Design web app (incl. feedback to respondent)	
#16	Design mobile app (incl. feedback to respondent)	
#17	Infrastructure logic web app	
#18	Infrastructure logic mobile app	
#19	Business logic web and mobile app	
Local storage e-HBS		
#20	Local storage device	
Availability application e-HBS		
#--	Web app infrastructure	✓
#--	Mobile app infrastructure	✓
#--	Updates applications	✓
Availability application e-HBS		
#21	Update Client API to database	✓
#22	Update Client API to plugins/microservices	✓

4.2 Communication of the results

The final results of the project were shown during the workshop (see D1.3).

Workshop info, agenda and PowerPoint presentations are available on the Eurostat wiki page: https://webgate.ec.europa.eu/fpfis/wikis/display/ISTLCS/Workshop2022.

The presentation of the CRŒSS project is also available on our Webpage ITNRW: "NRW connect extern" as B5460_BE_1.3_Workshop_Presentation_ENG.pdf

- The workshop presentation contains the slides presented by Statbel, hbits and Destatis, as well as a live demo about the new UI/UX.

During the course of the project, we have updated the leaflet with general information about the project that was made by Eurostat during the SOURCE™ project. The leaflet of the CRŒSS project is also available on our Webpage ITNRW: "NRW connect extern" as B5460_BE_1.3_Workshop_LEAFLET_ENG.pdf. We distributed this leaflet during the Task Force on innovative tools and sources for HBS/TUS the day after the workshop.

A webpage ITNRW: "NRW connect extern" is available with an overview of all the deliverable reports and links to external sources. A Username (E-mail) and Password are needed and van be provided by Destatis. The website is available in Deutsch and English. (see D1.5)

All presentations given during the coordination meetings and final workshop will be available in the dedicated sections, such as:

- WP1 – D1.3 – Workshop
- WP4 – D4.1 – UI/UX for TUS and HBS

Live demonstrations inbedded in presentations:

- WP3 – D3.2 – Geolocation microservice
- WP3 – D3.3 – Receipt scanning microservice prototype
- WP3 – D3.4 – R-cran package microservice
- WP4 – D4.1 – UI/UX for TUS and HBS
- WP4 – D4.2 – Report on the usability tests
- WP5 – D5.1 – Document database technology pillars

The final presentation and live demonstration given during the workshop (see D1.3) gives the latest overview of the developments done for the CRŒSS project.

The Github directory for code on WP3 – D3.2 – Geolocation microservice and WP3 – D3.3 – Receipt scanning microservice prototype was made available for Eurostat through ZIP-files who can distribute the code as they see fit. The same procedure was applied for the WP3 – D3.4 – R-cran package microservice.

5   WP1 – D1.5 – Assessment communication

Overall goal: webpage that communicates information about the project.

Webpage ITNRW: "NRW connect extern". A Username (E-mail) and Password are needed and van be provided by Destatis. The website is available in Deutsch and English.

Figure 12: Login screen website "NRW connect extern"

This website will contain the final report as well as all separate deliverables and attachments, such as the translated files.

Workshop info, agenda and PowerPoint presentations are available on the Eurostat wiki page: https://webgate.ec.europa.eu/fpfis/wikis/display/ISTLCS/Workshop2022.

An overview of the project is also given on the Eurostat website: https://ec.europa.eu/eurostat/web/household-budget-surveys/projects as well as on the Statbel website https://statbel.fgov.be/en/themes/households/household-budget-survey-hbs#documents.

WP2 – Definition of a cross-domain data collection platform

<u>State of play</u>

A cross-domain tool is beneficial for researchers and NSIs in an organisational perspective. To date the Inventory showed no tool that services both TUS and HBS.

<u>Added value</u>

The objective of WP2 was to upgrade the MOTUS-tool from a tool for TUS, to a tool that can service both TUS and HBS.

<u>Actions</u>

One of the main characteristics of MOTUS is its modularity, by means of the different research builders. These different builders defined for TUS (12 in total) can be reused for HBS. This makes the step-up to incorporate HBS into MOTUS more than feasible. To upgrade MOTUS to an HBS-platform the main task is to define a new diary builder for HBS.

<u>Objectives</u>

The specific objectives are linked to the enlarging of the back-office of MOTUS to include HBS, besides TUS. This project reuses the MOTUS components or so-called 'builders' that are already available. To include HBS the definition of a new diary builder (specifically to HBS) is important. The technical components will be available at the start of the project.

WP2 focused on the definition of the different underlying parameters of the new diary component and the reuse of the other components or builders to define a HBS-research.

The (iterative) testing of the input ran in accordance to WP4 'Defining and testing of new interfaces for TUS and HBS'. The specific objectives of WP2 were:

- Definition/evaluation of the back-office architecture of MOTUS to include HBS.
- Definition of HBS research components and research flow in MOTUS.
- Redefinition of the client API.

6 WP2 – D2.1 – Documentation research components

Overall goal: chapter final report.

Still, a complete and finalised version of questionnaires or flow is not available at the moment. Current discussions on HBS-contents (IESS), division among household/individual questionnaires and COICOP transmission makes it a continuing process. The very detailed documentation in SOURCE™-project was done based on the very detailed specifications of the HETUS-guidelines. For HBS and therefore the CRŒSS project, namely the delivery of 2.1 Research components it will be harder to come to a comparable detailed documentation.

Therefore, when reading this document, keep in mind that the following must be seen as intermediary and still changing and growing input. It will also live from ongoing discussions between the consortium of Statbel, Destatis and hbits.

6.1 General description on the components and modularity of MOTUS

As already documented in the SOURCE™-project for TUS, MOTUS is particular powerful because of the modularity of the components in the back-office. The available research components can be (re)defined within the MOTUS platform itself. At the same time, the platform can assure within a given context comparability, reliability and quality of the data.

In doing so, an NSI can design survey questionnaires, expense diaries and communication strategies. Next, all research elements can be combined into an automated flow. Once configured, the research flow runs automatically, including communication with respondents.

Compared to the very clear boundaries described in the HETUS-guidelines, specifications for HBS give NSIs more freedom of choice in ESS-comparable data that need to be collected. For components that are clearly set out, we will use the components of MOTUS to program the HBS-ecosystem as being known from the specifications of HBS. For the e-diary it means the setup of:

- a household survey (for reference persons),
- an individual survey (for all household members),
- an expense diary,
- context questions (e.g. discount, country).

Considering the outcomes from the questionnaire on tools and sources the different parameters of the HBS e-diary need to be considered:

- number of diary days / length of the observation period,
- household level/composition,
- place of expense,
- classification of expense (according COICOP classification).

In the following, we like to describe the tasks we have taken within the CRŒSS project to further define the different components that are relevant for an HBS study within MOTUS. Questionnaires can be built up towards HBS accordingly TUS. That means that the MOTUS builder can be used in the same way as for TUS-questionnaires. The diary - for TUS based on activities - received a complete new design to address new developments but also to line up with the expense-related diary. The new design is documented in a separated

delivery 12 (w4.1 UI/UX MOTUS, page 23 following). In the diary context the classification needs a shift towards the COICOP classification or related national classifications.

Household level/composition (My group function) -> see also deliverable 12 on WP4.1 on new UI/UX MOTUS.

More than in TUS, for HBS the household perspective as well as the individual perspective is important. MOTUS has developed a new My Group-component, which can be used in respect to the household budget survey in order to composite the household, 'My Household'.

Per household member the first name, last name, email, birthday and gender is asked to complete, as a first step. In a second step the relation to the reference person, and if more people are living in the household, to the other household members.

The relations to which the reference person can choose from can be defined in the back-office of MOTUS. The relation can be gender-related defined (male, female, x and generic). An example is brother, sister and sibling. For every position a further description can be given.

Equally important, also the reverse relation can be defined. An example is mother and daughter. When the relation between a mother towards her daughter is defined, MOTUS can then also display the reverse relation, being the relation between a daughter and her mother without having to fill it in twice. This is a value since no interviewer is available to make these translations and because reference persons will need the opportunity to edit the relations between the household members. As redefinition/adaptation is possible to be in a random order, all relations need to be visible, also the reverse relations.

The 'My Group'-component belongs to a research in particular, and so it can be defined modular by any NSI.

6.2 COICOP classification

In order to come to a harmonised list Destatis and Statbel concluded after several meetings to use the German (SEA-)classification (and the herein available extensive list of tags) and to translate these to the French, Dutch (FR, NL) as well as to English (EN). T

he classification chosen is still based on E-COICOP due to the ongoing transmission of the national classification towards UN-COICOP2018. The consortium has chosen to include a classification list to which at the most detailed level (7 levels) tags were available. This decision is also taken because of the interlinkage with the scan receipt microservice which benefits strongly from detailed information. Outside of the PoC the classification can be replaced through other NSI lists.

6.3 Research flow general study and questionnaires

One of the main characteristics of MOTUS is that the platform supports online data collection via automated processes. The presented research flow discusses the various stages the respondents have to go through in order to complete their participation successfully. The research flow also takes into account the various actions and communications that are part of every stage. The description below shows an exemplary research flow in MOTUS. For every research (as well as for every country) this flow can be different, as it is been built up by modular components.

The household budget survey is a research done on the household level. Therefore, a distinction is made between the reference person of the household and the other 'eligible' household members to participate to the research. The reference person has to complete

certain household related tasks, while the other household members are asked to complete individual tasks to sum up to the household level, and which are in sync with the household level. MOTUS can arrange in sync moments, as being moments through the survey that every eligible household members should be in the same state. An example is at the beginning of the consumption diary. However, it can placed at different other places in the flow.

To make the household perspective feasible, MOTUS can relate different researches to each other. In this way, reference persons answer the household related questionnaires but can also define the household composition, which, then, is used to invite and related household members within the second research.

Figure 13: Research 1 - Reference person
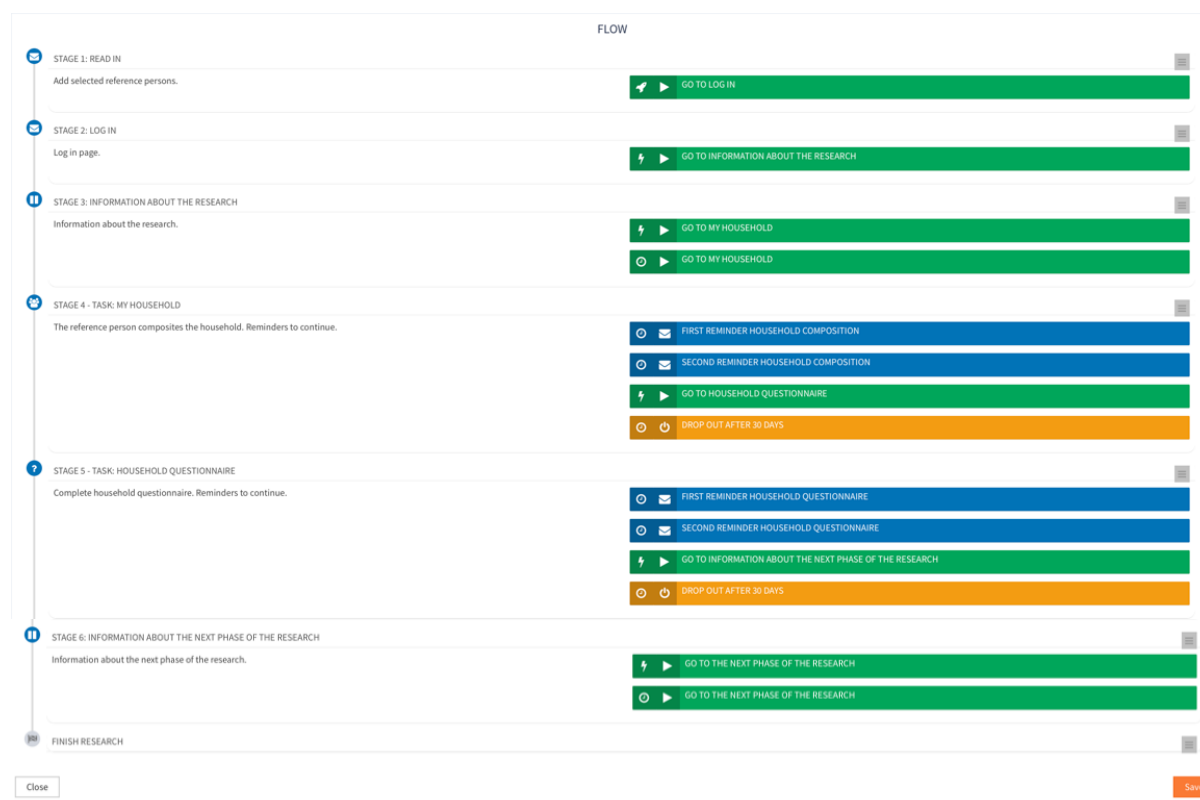
The research setup starts from the predisposition that respondents are invited to the study via a postal letter, as is the case for Statbel and Destatis at this moment. This invitation letter explains the research to the household, and provides a username and related password to take part in the study. These settings can be modified by other NSIs.

As the credentials are provided by postal letter, other people might have seen these initial credentials. Taking into account the GDPR rules MOTUS provides the option to include an obligatory password reset and activation of the account for every new user of the MOTUS platform. This action is always included at the first interaction of each 'user' with MOTUS. NSIs can choose whether to activate this process or not.

Reference persons will meet this action in research 1, while the other household members face this action in research 2.

Below an exemplary overview is given of the different stages in both of the researches. The parameters given in the overview are adaptable by each NSI. For Germany time marks for reminders need to be adapted, as respondents need to keep the diary for 3 months.

Figure 14: Research 2 – Household members
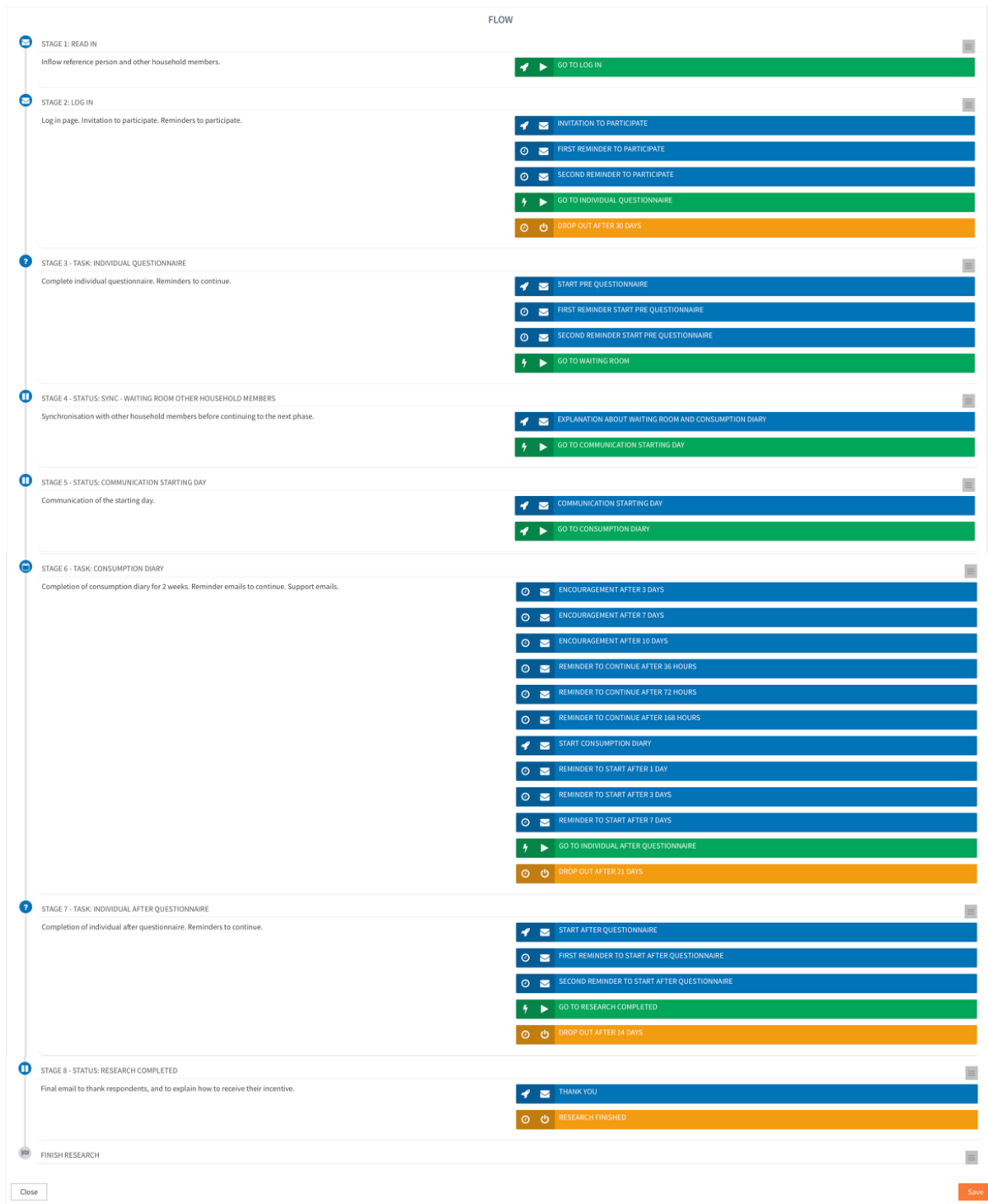
Table 5: Example of different stages of the research - reference person

Research 1 – Reference person
Stage 1: Read in
In order to invite the reference person of the household to participate to the study, MOTUS needs to know the contact details of the respondent, and more in particularly the email address. In the best-case scenario also other background information is provided to MOTUS. There are two options: one is to upload a file of respondents to MOTUS, another is to enter the respondents via an API-link. This API-link can as an example be linked to a webpage where respondents can subscribe themselves to this study.
Action: Once reference persons are linked to the study they will automatically go to the next stage.
Stage 2: Log in
The reference person logs into the study. If this is not done before predefined moments Statbel and Destatis resends a postal letter to remind the household for participation.
Action: Go to the information page about the research
Stage 3: Information about the research
The reference person is given information about the (setup of the) research.
Action: Go to My Household by confirming the read of the page
Action: Go to My Household after waiting to the next day until 10:00
Stage 4 – Task: My Household
The reference person composites the household in accordance to the HETUS-guidelines.
Communication: First reminder to continue on day 1 at 10:00
Communication: Second reminder to continue on day 3 at 14:00
Action: Go to Household questionnaire at completion
Action: Drop out after 30 days
Stage 5 – Task: Household questionnaire
The completion of the household questionnaire. The household questionnaire is based on the HETUS-guidelines. See additional document B5460_BE_HBS_Household_Questionnaire_language.xlsx.
Communication: First reminder to continue on day 1 at 10:00
Communication: Second reminder to continue on day 3 at 14:00
Action: Go to Information about the next research at completion
Action: Drop out after 30 days
Stage 6 – Information about the next phase of the research
The reference person is given information about the next phase of the research.
Action: Go to the next research by confirming the read of the page.
Action: Go to the next research after waiting to the next day until 10:00.
Finish research
The reference persons are linked to the next research.
At this moment, also the household members are invited to participate to research 2.
The reference person of the household receives an email (specified in the 'My Group'-component) with the credentials of all 'eligible' household members to participate to the second research. In HETUS the age of the respondent marks the invitation to participate. This is set on 10 years old, but can be modified by an NSI. For Destatis in HBS this will be the age of 16. Also, the individual members receive a person email with (only) their personal credentials. The email to the reference person with all credentials assures to have a confirmed email address. Just as in the previous research, the first action for new users is to change their password, and to confirm their account. At that moment, the credentials in the email to the reference person become invalid.
Action: All in organized automatically in the 'My Group'-component.

Table 6: Example of different stages of the research - household members

Research 2 – Household members
Stage 1: Read in
In order to invite all 'eligible' household members to participate to the study. These household members flow in from the previous linked research (research 1).
Action: Go to Log in immediately
Stage 2: Log in
Log in to MOTUS.
Communication: Invitation to participate, immediately
Communication: First reminder to continue on day 1 at 10:00
Communication: Second reminder to continue on day 3 at 14:00
Action: Go to Individual questionnaire on completion
Action: Drop out after 30 days
Stage 3 – Task : Individual pre questionnaire
The completion of the individual pre questionnaire. The individual questionnaire is based on the HETUS-guidelines. See additional document B5460_BE_HBS_Individual_Questionnaire_language.xlsx.
Communication: Email at the start of the pre questionnaire, immediately
Communication: Second reminder to continue on day 3 at 14:00
Action: Go to Information about the next research at completion
Stage 4 – Status : Sync – Waiting room other household members
Synchronization with other household members before continuing to the next phase.
Communication: Explanation about waiting room and consumption diary
Action: Go to Communication starting day, immediately
Stage 5 – Status : Communication starting day
Communication of the starting day.
Communication: Communication starting day, immediately
Action: Go to Consumption diary at Defined date
Stage 6 - Status: Consumption diary
The completion of a Consumption diary, based on the COICOP classification list, and the contextual variables to the COICOP list. See additional document B5460_BE_HBS_Context_questions_language.xlsx
Communication: Start consumption diary, immediately
Communication: Reminder to start after 1 day
Communication: Reminder to start after 3 days
Communication: Reminder to start after 7 days
Communication: Encouragement after 3 days
Communication: Encouragement after 7 days
Communication: Encouragement after 10 days
Communication: Reminder to continue when no logs for 36 hours
Communication: Reminder to continue when no logs for 72 hours
Communication: Reminder to continue when no logs for 168 hours
Action: Go to Individual questionnaire
Action: Drop out after 21 days
Stage 7 – Task: Individual after questionnaire
The completion of the individual after questionnaire.
Communication: Email at the start of the after questionnaire, immediately
Communication: First reminder to continue on day 1 at 10:00
Communication: Second reminder to continue on day 3 at 14:00
Action: Go to Research completed at completion
Action: Drop out after 14 days

Stage 8 – Status: Research completed
Final email to thank respondents, and to explain how to receive their incentive.
Communication: Thank you, immediately
Action: Research finished, next day at 12:00

6.4 Available languages

The research will have as a default the English language and as secondary languages German, French and Dutch. If the language preference is known by MOTUS, the respondent receives the research and communication in his/her preferred language.

6.5 Overview of the (translated) documents

All documents are Excel documents. The starting documents were in German. The translations towards French, Dutch and English were done by the translation department of Statistics Belgium.

For the ECOICOP, the German SEA classification was used. The "tags" were translated to English, French and Dutch, but the COICOP aggregation was done from the fifth level and up by the official ECOICOP nomenclature from the RAMON database. Not all tags were translated from German. This was not possible, because some vocabulary is country-specific and was very burdensome.

- B5460_BE_2.1_HBS_ECOICOP_SEA_GER.xlsx
- B5460_BE_2.1_HBS_ECOICOP_SEA_TAGS_GER.xlsx
- B5460_BE_2.1_HBS_ECOICOP_SEA_TAGS_ENG.xlsx
- B5460_BE_2.1_HBS_ECOICOP_SEA_TAGS_FR.xlsx
- B5460_BE_2.1_HBS_ECOICOP_SEA_TAGS_NL.xlsx

Destatis has divided the national stores into categories (see D4.1). This division has been translated to English: B5460_BE_2.1_UI-UX_HBS_Shoplist_ENG.php

Statbel had a register that was used for the Consumer Price Index until 2013. At a certain point, this list was abandoned. A provisional list was constructed based on NACE code '47' and can be found as the following document: B5460_BE_2.1_UI-UX_HBS_Shoplist_NL.xlsx

Destatis developed an individual, household and context questionnaire, based on the questionnaires developed in the SOURCE™ project.

- B5460_BE_2.1_HBS_Context_Questions_GER.xlsx
- B5460_BE_2.1_HBS_Household_Questionnaire_GER.xlsx
- B5460_BE_2.1_HBS_Individual_Questionnaire_GER.xlsx

The MOTUS UI/UX was developed in English by hbits/VUB and translated to Dutch and French by the translation department of Statistics Belgium in .xliff files. This way hbits/VUB had the possibility to upload the translated content immediately into the UI/UX.

- B5460_BE_2.1_MOTUS_FR.xliff
- B5460_BE_2.1_MOTUS_NL.xliff
- B5460_BE_2.1_MOTUS_ENG-NL.xliff
- There is also an overview in html available: B5460_BE_2.1_MOTUS_ENG-NL.html

7 WP2 – D2.2 – Technical documentation DB and API

Overall goal: technical documentation.

In order for MOTUS to become a cross-domain data collection platform, a new HBS diary component or builder needs to be introduced. This document describes the required database changes and the client API modifications to support such an HBS component.

7.1 HBS Database technical documentation

7.1.1 Overview

Conceptually the HBS database consists of 2 parts:

- respondent's expenses: when, where and which purchases a respondent did

  COICOP: organization of the COICOP classification including fixed costs, tags for searching etc.

Data confirmed by the respondent is stored 'as is' in the database so that no information can be lost due to intermediate processing. Data processing (analytics) is not immediately performed.

There is a link possible to externals databases which contain store and product information . Those external databases are not part of MOTUS, they can used as external sources of (additional) information. This information could be obtained by scraping store websites, retrieving payment information from financial institutions, connecting to store databases etc.

Figure 15: HBS database technical documentation - overview
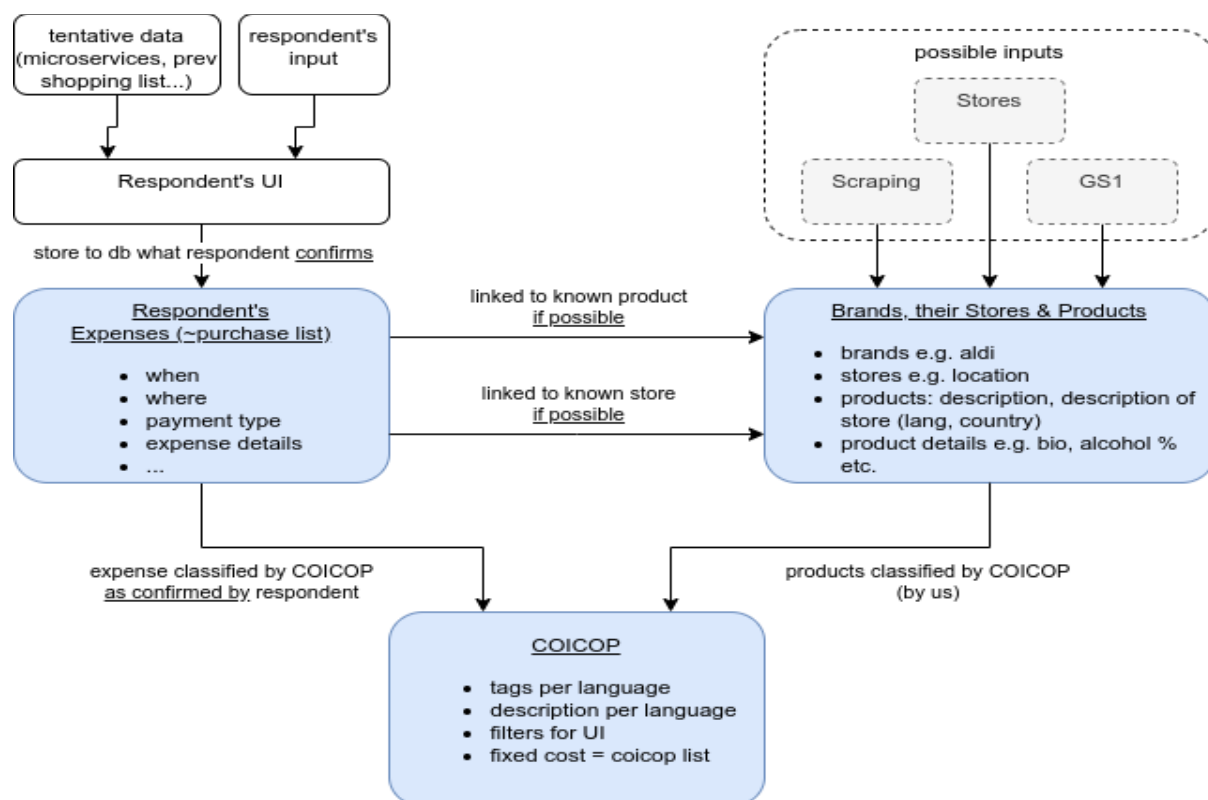
7.2 HBS database structure

Respondents can input two types of expenses: purchases which are typically but not exclusively done in a store (table mts_expenses) and fixed costs such as renting or energy costs (mts_fixed_costs).

Expenses consist of expense items. Each expense item is linked to a COICOP code. There is also a table for discounts at expense level (mts_expenses_discounts).

The fixed costs fall into a set of categories which can also be linked to COICOP codes.

The currency is not taken into account yet.

7.2.1 mts_expenses

This table contains:

- when (date, time, timezone),
- where (country and store description, see later),
- what kind of expense (professional – and how much of it is for private use), and
- how it is paid (credit card or not).

7.2.2 mts_expenses_discounts

A single discount. The 'slug' indicates which kind of discount: e.g. discount coupon, customer discount, one time action etc.

7.2.3 mts_expense_items

The items of a expense. The table includes:

- a description of the item (as given and confirmed by the respondent),
- whether it is a return item or not,
- the total price (calculated from quantity*price - discount),
- number of items (e.g. 6 bottles in a pack),
- quantity_unit of an item (e.g. 33cl),
- unit e.g. cl, kg etc.,
- a discount variable,
- whether to share it with household or not (keep_private),
- link to COICOP code.

7.2.4 mts_fixed_costs

Fixed costs are similar to recurring expenses e.g. energy, housing etc. It is therefore similar to an expense but is linked to a fixed cost category and it has a frequency (e.g. every month).

7.2.5 mts_fixed_cost_categories

The 'slug' describes the category e.g. mobility, telecommunications. It is hierarchical e.g. leasing a car is a part of mobility, and it can have a set of COICOP codes.

Figure 16: HBS database structure overview – part 1
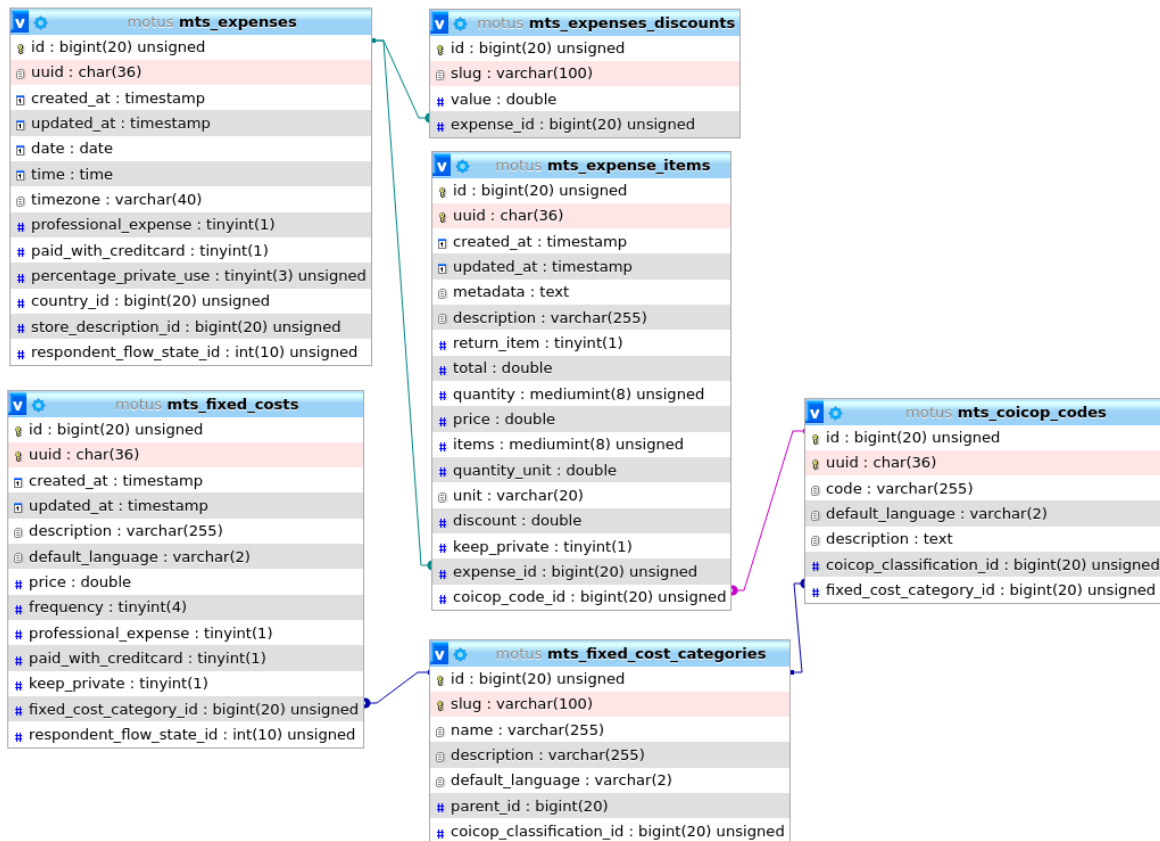
In the UI a respondent must give a description of where an expense has been made. That description is stored in the store description table, together with the store categories it belongs to.

7.2.6 mts_store_descriptions

A store description contains the description given by a respondent. The respondent id is also part of the table so previously visited locations can be easily found. Metadata is a json field which might contain details of an external store on which the description is based.

7.2.7 mts_store_categories

The different store categories. Examples are supermarket, local market, online, restaurant etc. The slug is used instead of an idea for communication with front-end.

7.2.8   mts_store_category_external_store

Example table of how MOTUS would map external stores to internal store categories. The assumption is that external database cannot be changed.

Figure 17: HBS database structure overview – part 2
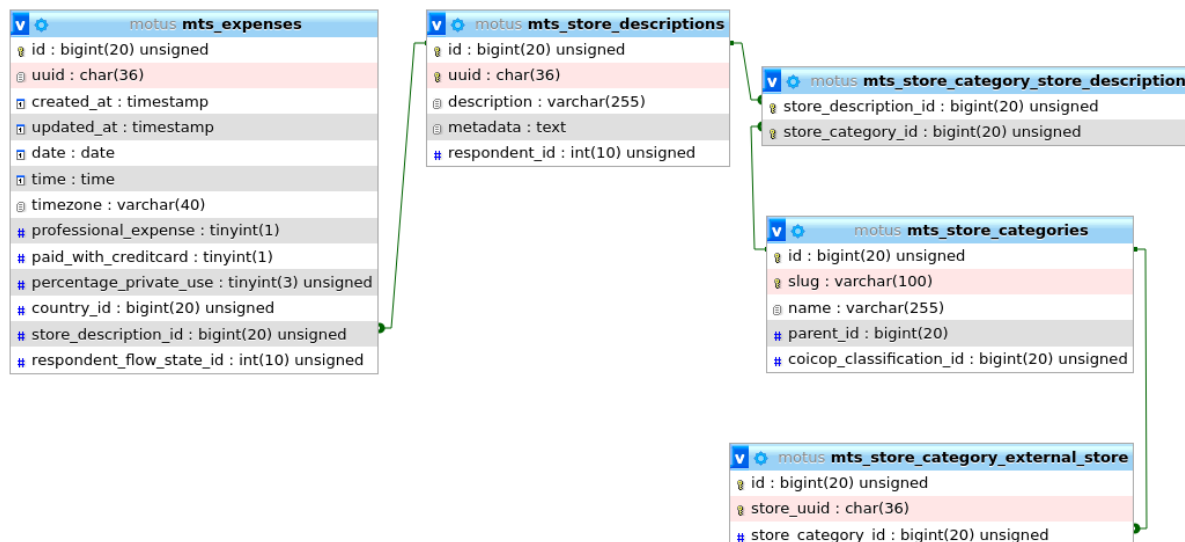
7.3   COICOP classification

7.3.1   mts_COICOP_classifications

The top of the COICOP-table hierarchy is mts_coiciop_classificiations. This table allows MOTUS to support different COICOP classifications. The format property of the classification describes the format of a COICOP code e.g. additional letters/digits might be present beyond the European 4 level hierarchy.

7.3.2   mts_COICOP_filters

Used in the front-end UI to filter the COICOP search results.

7.3.3   mts_COICOP_codes

Code and description of a COICOP code. The default language field is used for translations.

7.3.4   mts_COICOP_code_tags

Tags per language (and per classification) to find COICOP codes.

Figure 18: HBS database structure overview – part 3
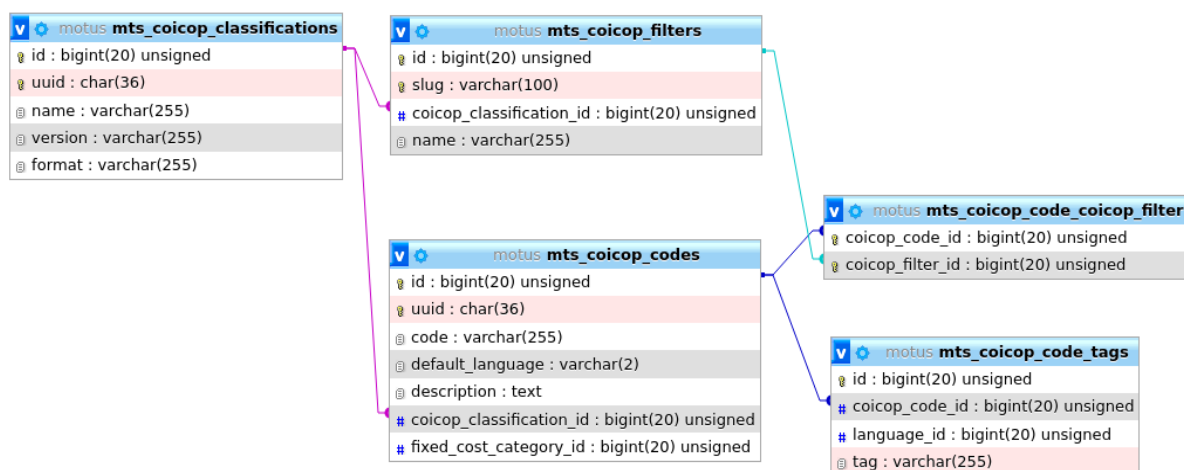
7.4 Other extension to the MOTUS database

Since the HBS component is a special type of diary, a few extensions were needed to the current diary concept. More specifically,

- barcode support Boolean,
- ticketscan_support Boolean,
- optional_expense_item_info: array of fields to show in frontend UI,
- default country and prioritized countries: list of countries to show first when booking an expenses.

7.5 API technical documentation

MOTUS exposes a specific API to each component of the system (see Diagram below). This section focuses on the Client API which is the interface between the MOTUS front office (web or mobile app) and the MOTUS backend.

Figure 19: API for each component in the system

STATBEL
Belgium in figures

D STATIS
wissen.nutzen.

VUB VRIJE UNIVERSITEIT BRUSSEL

h

eurostat
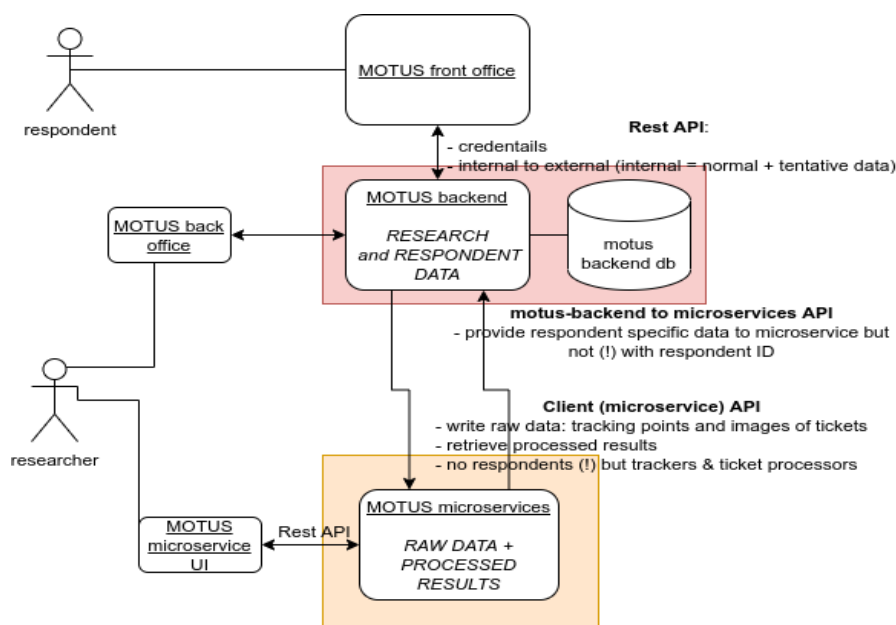
7.6 API endpoints

7.6.1 Day

Post

As a POST it is used to indicate that a specific day has no expenses (no_expenses_today).

Endpoint: /api/v1/diaries/{{ _.flowstate }}/day

Request Json:

```
{
        "date": "2022-03-18",
        "no_expenses_today": true,
        "expenses": null
}
```

Get

Retrieve all details of a specific day incl. Expenses.

Endpoint: /api/v1/diaries/{{ _.flowstate }}/day?date=2021-09-18

Response Json:

```
{
        "data": {
                "no_expenses_today": false,
                "expenses": [ <ARRAY OF EXPENSES – see below>
                ],
                "date": "2021-09-18"
        },
        "error": false
}
```

7.6.2 Expenses

Post

Endpoint: /api/v1/diaries/{{ _.flowstate }}/expenses

Request Json:

```
{
        "uuid": null,
        "date": "2021-09-18",
        "time": "10:00:00",
        "timezone": "Europe/Brussels",
        "professional_expense": true,
        "paid_with_creditcard": false,
        "percentage_private_use": 10,
        "total": 33,
        "country_code": "be",
        "store_description_uuid": "5a34f030-f06e-4e48-a4b4-5eff7633728c",
        "expense_items": [ <ARRAY OF EXPENSE ITEMS – see below>],
        "discounts": [
                {"slug":"discount_coupon","value":10.5}
        ]}
```

Get

Get all expenses: /api/v1/diaries/{{ _.flowstate }}/expenses

Get specific expense: /api/v1/diaries/{{ _.flowstate }}/expenses/2f2055fc-1c81-459a-983c-2308afc85175

Response Json:

```
{
        "data": [
                {
                        "uuid": "5014f00c-9a15-4002-9dca-89a93f9d204c",
                        "date": "2022-03-27",
                        "time": "12:00:00",
                        "timezone": "Europe\/Brussels",
                        "professional_expense": true,
                        "paid_with_creditcard": false,
                        "percentage_private_use": 69,
                        "total": 41.4,
                        "country_code": "be",
                        "store_description": {
                                "uuid": "6047adfb-f434-44cb-82da-debaf0de38d7",
                                "description": "Carrefour Market",
                                "metadata": {
                                        "store_uuid": "b817d90e-8e47-11ec-b1b5-
002590f96830"
                                },
                                "categories": [
                                        "stores-supermarket"
                                ],
                                "origin": null
                        },
                        "expense_items": [
                                {
                                        <ARRAY OF EXPENSE ITEMS – see below>
                                }
                        ],
                        "discounts": [
                                {
                                        "slug": "customer_discount",
                                        "value": 1.3
                                }
                        ],
                        "created_at": "2022-03-07T11:04:49.000000Z",
                        "updated_at": "2022-03-27T09:09:29.000000Z"
                }
        ],
        "error": false
}
```

Put

Same as POST, except the "uuid" should reference an existing expense.

Del

Deletes an expense.

Endpoint: /api/v1/diaries/{{ _.flowstate }}/expenses/119329fc-6374-492d-8979-b260adbb873b

7.6.3 ExpenseItems

Post

Creates a new expense item on an existing expense.

Endpoint: /api/v1/diaries/{{ _.flowstate }}/expenses/2f2055fc-1c81-459a-983c-2308afc85175/items

Request Json:

```
{
        "uuid": null,
        "metadata": null,
        "description": "expense item",
        "return_item": false,
        "price": 3.14,
        "quantity": 4,
        "items": 2,
        "quantity_unit": 500,
        "unit": "gr",
        "discount": 3.2,
        "keep_private": true,
        "COICOP_code_uuid": "df6badb6-ed09-4d5e-b33d-900f918f171e"
}
```

Get

Get all: /api/v1/diaries/{{ _.flowstate }}/expenses/2f2055fc-1c81-459a-983c-2308afc85175/items

Response Json:

```
{
        "data": [
                {
                        "uuid": "cbb6644e-11ce-48b5-9cb5-534aec799af3",
                        "metadata": null,
                        "description": "expense item 1",
                        "return_item": false,
                        "total": 24.34,
                        "price": 3.14,
                        "quantity": 11,
                        "items": 1,
                        "quantity_unit": 500,
                        "unit": "gr",
                        "discount": 10.2,
                        "keep_private": true,
                        "COICOP_code": {
                                "uuid": "df6badb6-ed09-4d5e-b33d-900f918f171e",
                                "code": "0111101",
                                "description": "Reis, z. B. im Kochbeutel"
                        },
                        "created_at": "2021-12-15T09:43:54.000000Z",
                        "updated_at": "2021-12-15T09:43:54.000000Z"
                }
        ],
        "error": false
}
```

Get specific one: /api/v1/diaries/{{ _.flowstate }}/expenses/3d4a2918-6ce1-4dd1-a83f-5738dee464d4/items/cbb6644e-11ce-48b5-9cb5-534aec799af3

Response Json:

```
{
        "data": {
                "uuid": "cbb6644e-11ce-48b5-9cb5-534aec799af3",
                "metadata": null,
                "description": "expense item 1",
                "return_item": false,
                "total": 24.34,
                "price": 3.14,
                "quantity": 11,
                "items": 1,
                "quantity_unit": 500,
                "unit": "gr",
                "discount": 10.2,
                "keep_private": true,
                "COICOP_code": {
                        "uuid": "df6badb6-ed09-4d5e-b33d-900f918f171e",
                        "code": "0111101",
                        "description": "Reis, z. B. im Kochbeutel"
                },
                "created_at": "2021-12-15T09:43:54.000000Z",
                "updated_at": "2021-12-15T09:43:54.000000Z"
        },
        "error": false
}
```

Put

Same as POST, except the "uuid" should reference an existing expense item.

Del

Deletes an expense item.

Endpoint: /api/v1/diaries/{{ _.flowstate }}/expenses/2f2055fc-1c81-459a-983c-2308afc85175/items/e3342171-528c-4232-8463-89f3bfef3251

7.6.4   FixedCosts

Post

Endpoint: /api/v1/diaries/{{ _.flowstate }}/fixedcosts

Request json:

```
{
        "uuid": null,
        "description": "fixed cost",
        "price": 3.14,
        "keep_private": true,
        "professional_expense": false,
        "paid_with_creditcard": true,
        "frequency": "annually",
        "fixed_cost_category": "rent_and_charges"
}
```

Get

Get all fixed costs: /api/v1/diaries/{{ _.flowstate }}/fixedcosts

Response json:

```json
{
        "data": [
                {
                        "uuid": "d0d7c03d-ccc6-4835-a314-72ed5f2d313e",
                        "description": "fixed cost",
                        "professional_expense": false,
                        "paid_with_creditcard": true,
                        "keep_private": true,
                        "price": 3.14,
                        "frequency": "annually",
                        "fixed_cost_category": "rent_and_charges",
                        "created_at": "2022-03-29T13:19:22.000000Z",
                        "updated_at": "2022-03-29T13:19:22.000000Z"
                }
        ],
        "error": false
}
```

Get specific one: /api/v1/diaries/{{ _.flowstate }}/fixedcosts/d0d7c03d-ccc6-4835-a314-72ed5f2d313e

Response json:

```json
{
        "data": {
                "uuid": "d0d7c03d-ccc6-4835-a314-72ed5f2d313e",
                "description": "fixed cost",
                "professional_expense": false,
                "paid_with_creditcard": true,
                "keep_private": true,
                "price": 3.14,
                "frequency": "annually",
                "fixed_cost_category": "rent_and_charges",
                "created_at": "2022-03-29T13:19:22.000000Z",
                "updated_at": "2022-03-29T13:19:22.000000Z"
        },
        "error": false
}
```

Put

Same as POST, except the "uuid" should reference an existing expense.

Del

Endpoint: /api/v1/diaries/{{ _.flowstate }}/fixedcosts/ddb94c45-65fd-4377-a0ff-e92dc2c7624e

7.6.5 FixedCostCategories

Get

Endpoint: /api/v1/diaries/{{ _.flowstate }}/fixedcostcategories

Response json:

```
{
        "data": [
                {
                                "slug": "rent_and_charges",
                                "name": "rent and charges",
                                "description": "rent, common costs, services...",
                                "subcategories": [
                                        {
                                                "slug": "rent_and_charges-
rent_of_property_without_charges",
                                                "name": "rent of property without charges",
                                                "description": null,
                                                "subcategories": []
                                        },
                                        {
                                                "slug": "rent_and_charges-
rent_of_property_with_charges",

                                                "name": "rent of property with charges",
                                                "description": null,
                                                "subcategories": []
                                        },
...
]}
```

7.6.6   StoreDescriptions

Post

Endpoint: /api/v1/storedescriptions

Request json:

```
{
        "uuid": null,
        "description": "my store description 2",
        "categories": ["stores-supermarket","stores-local_market"],
        "metadata": null
}
```

Get

Get all of a respondent: /api/v1/storedescriptions

Response json:

```
{
        "data": [
                {
                        "uuid": "89daff25-02a1-49a4-ad54-4d09274904ed",
                        "description": "Carr",
                        "metadata": null,
                        "categories": [
                                "stores-supermarket"
                        ],
                        "origin": null
                },
                {
```

```
                "uuid": "6047adfb-f434-44cb-82da-debaf0de38d7",
                "description": "Carrefour Market",
                "metadata": {
                        "store_uuid": "b817d90e-8e47-11ec-b1b5-002590f96830"
                },
                "categories": [
                        "stores-supermarket"
                ],
                "origin": null
        }
    ],
    "error": false
}
```

Get specific one: /api/v1/storedescriptions/5a34f030-f06e-4e48-a4b4-5eff7633728c

Response Json:

```
{
        "data": {
                "uuid": "5a34f030-f06e-4e48-a4b4-5eff7633728c",
                "description": "my store description 2",
                "metadata": null,
                "categories": [
                        "stores-supermarket",
                        "stores-local_market"
                ],
                "origin": null
        },
        "error": false
}
```

Search a store: /api/v1/storedescriptions/storesearch?q=ahn%20G&country=be&limit=10

Response json:

```
{
        "data": [
                {
                        "uuid": null,
                        "description": "Hahn Group",
                        "metadata": {
                                "store_uuid": "66542865-b61f-4606-936e-a66c1b1b1dcb"
                        },
                        "categories": null,
                        "origin": null
                }
        ],
        "error": false
}
```

Put

Same as POST, except the "uuid" should reference an existing store description.

Del

Deletes a store description.

Endpoint: /api/v1/storedescriptions/2f45b211-ee53-447d-bd22-e20ae1d4b4da

STATBEL
Belgium in figures
DESTATIS
wissen.nutzen.
VUB VRIJE
UNIVERSITEIT
BRUSSEL
h
eurostat

7.7 COICOP

Get

Endpoint:
api/v1/diaries/{{_.flowstate}}/COICOPcodes?q=reis%2Ckartoff&filters=foodstuffs&limit=20

Response json:

```json
{
        "data": [
                {
                        "matching_tags": [
                                "Reis",
                                "Basmatireis",
                                "Bruchreis",
                                "Expressreis",
                                "Kochbeutelreis",
                                "Langkornreis",
                                "Milchreis",
                                "Naturreis",
                                "Paella-Reis",
                                "Reis geschliffen",
                                "Risotto-Reis",
                                "Sella Reis",
                                "Sushi-Reis",
                                "Suppenreis",
                                "Wildreis"
                        ],
                        "COICOP_code": {
                                "uuid": "df6badb6-ed09-4d5e-b33d-900f918f171e",
                                "code": "0111101",
                                "description": "Reis, z. B. im Kochbeutel"
                        }
                },
                {
                        "tags": [
                                "Djuvec Reis",
                                "Gemüsereis",
                                "Reis-Pyramiden",
                                "Reissalat",
                                "Reiszubereitungen"
                        ],
                        "uuid": "3bbc154b-7d55-412f-9949-815c4c9244fb",
                        "code": "0111109",
                        "description": "Reiszubereitungen mit Fleisch, Fisch, Meeresfrüchten
oder Gemüse"
                }
        ],
        "error": false
}
```

7.7.1 Overview

Get

Returns the diary overview.

Endpoint: /api/v1/diaries/{{ _.flowstate }}/overview

Response Json:

```json
{
        "data": {
                "expenses_total": 41.4,
                "days_in_diary": 27,
                "fixed_costs_monthly": 0.26
        },
        "error": false
}
```

7.7.2   Stats

Get

Endpoint: /api/v1/diaries/{{ _.flowstate }}/stats

Response Json:

```json
{
        "data": {
                "percentage_per_category":
"{\"01\":100,\"02\":0,\"03\":0,\"04\":0,\"05\":0,\"06\":0,\"07\":0,\"08\":0,\"09\":0,\"10\":0,\"11\":0,\"12\":0,\"13\":0,\"14\":0,\"15\":0}",
                "most_bought_product": "Pudding",
                "most_visited_store": "Carrefour Market"
        },
        "error": false
}
```

WP3 – Development of new TUS & HBS components (microservices)

> **State of play**
>
> Smart data can provide more detailed/accurate information on activities and purchases. For TUS and HBS some prototypes become available but need to be evaluated and do not (yet) have the status of a microservice.
>
> **Added value**
>
> By including microservices, respondents, researchers and NSIs become better supported. Microservices can gather smart data in favour of TUS and HBS. These microservices enlarge the capacity and functionality of MOTUS while they are also equally shareable to other platforms and purposes.
>
> **Actions**
>
> This project developed a software architecture incorporating microservices that is linked to the MOTUS-data collection platform, but which is equally reusable and shareable to other platforms by means of the APIs. The project aims to include geolocation, receipt scanning and R-cran packages as microservices.
>
> The connection with microservices is a long-term strategy. Different microservices can be attached to support the TUS and HBS data collection for respondents, researchers and NSIs. It is also a profound way to connect wearables or other smart devices in the future.
>
> **Objectives**
>
> In WP3 three new functionalities (microservices) are included. Two of the new functionalities integrate smart data to enrich the primary data and/or support the respondent in the data collection process. Through smart devices information can be collected which is otherwise more difficult to include.
>
> The first is the geolocation for TUS, the second is the receipt scanning for HBS. A third functionality (microservice) is the developed R-cran package for TUS and HBS. Through R-scripts various implementations can be developed to support the data collection and cleaning (quality checks, control) of the data. To include these microservices a shareable software architecture is needed, also comprising a database and API.
>
> The specific objectives of WP3 are:
>
> - Definition and development of a software architecture for microservices.
> - Development and integration of microservices to support the data collection, for TUS and HBS.

8 WP3 – D3.1 – Microservices for TUS and HBS

Overall goal: report describing the software architecture, database, API and inclusion of the 3 microservices.

8.1 Introduction

This document describes the software architecture of the MOTUS platform including microservices.

8.2 MOTUS architectural overview

MOTUS conceptually has 3 layers:

1. a database layer (MOTUS db in the Figure below). MOTUS requires a standard-SQL relational database.
2. an application layer: back-office which is mainly for researches, and the core or backend which contains the application logic for respondents
3. a presentation layer: the user interfaces. For the respondent this includes the web application and the mobile app (IOS, Android). For the researches this is the backoffice UI.

Figure 20: MOTUS architectural overview: 3-tiers
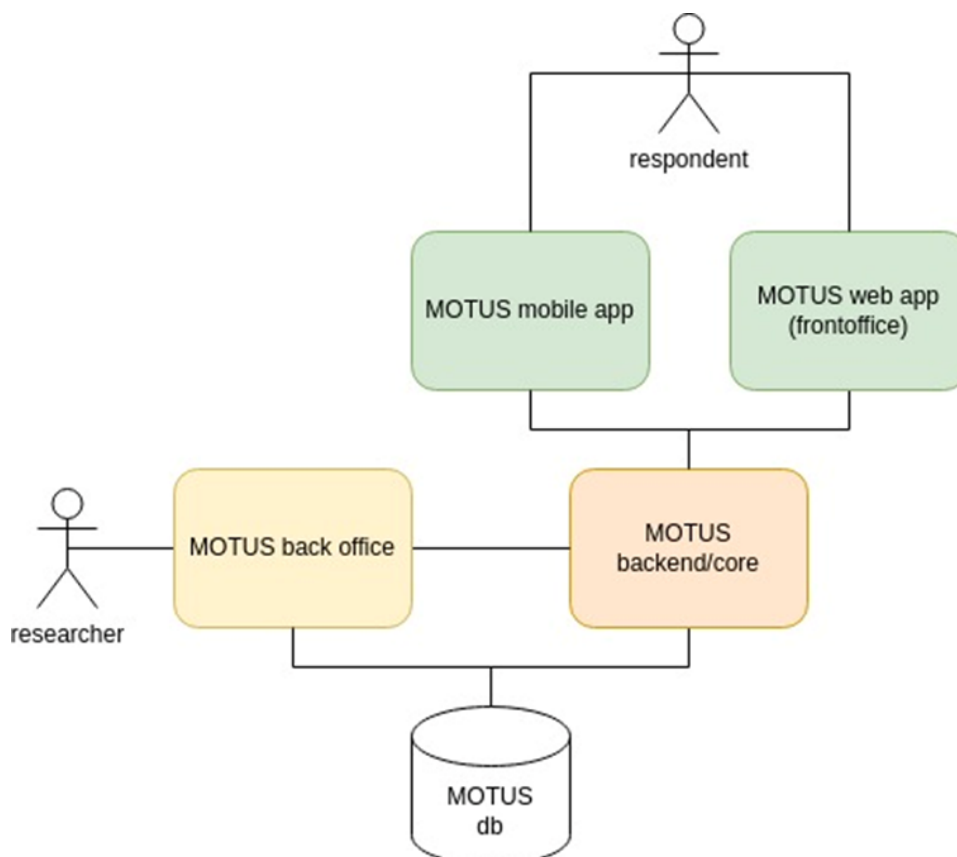
Given those 3 conceptual layers and their components, it is possible to apply a mapping to a physical architecture. Examples of physical architectures are:

- a 3-tier architecture in which each layer corresponds to a physical machines, or
- a cluster-based approach in which a collection of one or more physical machines is used for deployment.

The MOTUS platform can be extended with microservices as is explained in the next section. The GeoService and ReceiptScanner follow the generic microservice architecture, while R has a different architecture.

8.3 Generic microservice architecture

8.3.1 Introduction

The structure of this section is based on views and perspectives. Views illustrate the structural aspects of an architecture (e.g. where is data stored?), while perspectives consider the quality properties (e.g. scalability) of the architecture across a number of views.

Although in most diagrams the MOTUS core is used to illustrate interaction with the microservice, any other application could use the microservice for data processing tasks. The microservice is independent and not coupled to the MOTUS core.

8.3.2 Views

Context view

Next diagram shows how the MOTUS microservices fit into the general MOTUS architecture as a black box. A connection line indicates a relationship, dependency or interaction between 2 elements (a user, a system...).

Figure 21: Microservice architecture - context view
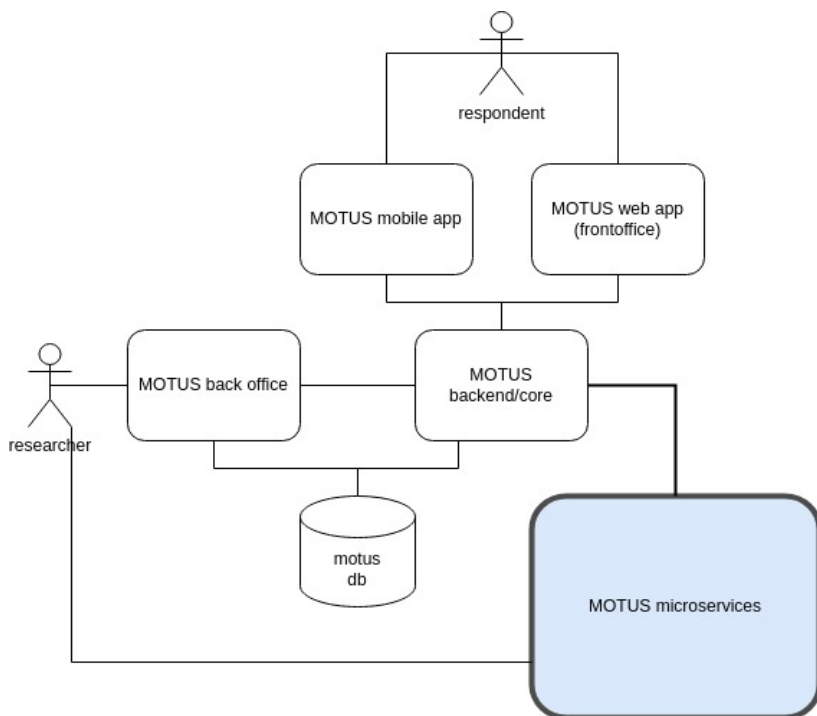
There are 2 types of users: the respondent and the researcher.

1. Respondents only interact with MOTUS via the front office. They are unaware of the existence of microservices.
2. Researcher typically interact with MOTUS via the back office. For some specific tasks (e.g. adding a global geolocation, investigating a receipt scan) they might want to use the microservice GUI(s).

It is important to mention that there is no direct link between MOTUS microservices and the MOTUS backend db. This means that the MOTUS backend has full control over which data is delivered to the microservices. The exact mechanism which guarantees privacy will be explained in the "Regulation perspective" section.

Functional view

This diagram describes the different runtime functional elements of the microservice.

Figure 22: Microservice architecture - functional view
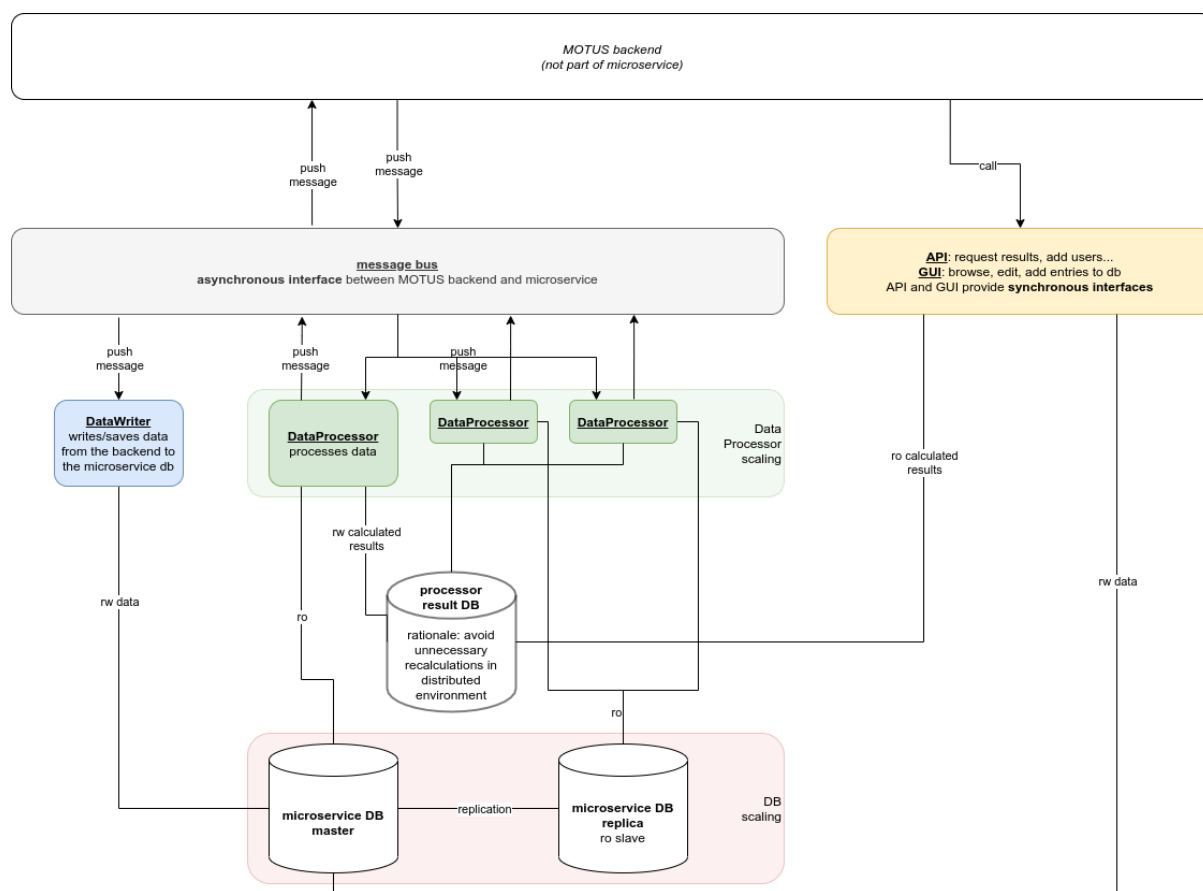
Main elements and their responsibilities are:

Message bus

The message bus allows for asynchronous communication between the MOTUS backend and the microservice.

Rationale:

- avoid blocking calls e.g. the motus backend must be able to quickly forward data (scanned receipt, tracking point) to the microservice without being blocked for too long. The message bus is able to fulfill this requirement by putting the data in a queue without any processing.

- notification service e.g. the DataProcessor can send a message that (some) data is processed. The motus backend can then take appropriate action.

Chosen technology: RabbitMQ (https://www.rabbitmq.com/).

API and GUI

Is the synchronous interface of the microservice.

The API is used by the MOTUS backend to fetch microservice data, request processed data results etc.

The GUI is used by a researcher or operator to:

- browse and inspect the results of the DataProcessor in the processor result DB,
- browse, inspect and edit the data of the microservice DB,
- possibly other functions e.g. add a scanned receipt and test the outcome.

DataWriter

Receives push messages with data from the motus backend via the message bus. The DataWriter writes the data in the microservice DB. Data push messages are queued in the message bus until the DataWrites is able to accept them.

There is only one DataWriter process in order to guarantee that the received data is written to the database in the same order as the data was pushed by the motus backend. This avoids (subtle) race conditions in which a DataProcessor start processing data with missing in-between data (e.g. a tracking point is missing in the db between the first and the last tracking point).

DataProcessor

Does the real work of processing the data.

Starts processing when it receives a push message of what to process. Note that a single push message is delivered to only and only one DataProcessor. Scalability is achieved by load balancing the requests over the DataProcessors. See the perspective on performance and scalability.

A DataProcessor pushes a message on the bus when processing is done.

Processor Result DB

This DB stores the results of the DataProcessors. Since results can be re-calculated, persistent storage is not a strict requirement e.g. one might opt to make it a RAM only db.

It is accessed by the API/GUI element to fetch the results.

It can be consulted by DataProcessors to avoid the re-calculating of data, it therefore also act as a cache.

Chosen technology: redis (https://redis.io/). It is also possible the replicate a redis db over several nodes if needed so (see perspective on performance and scalability).

Microservice DB

Used by the DataWriter to store pushed data.

Data is consulted by the DataProcessors.

GUI is able to change data if needed so.

DB scaling is achieved by replication, see perspective on performance and scalability perspective.

Information view

This view describes the way that the microservice stores, manipulates, manages and distributes information. The diagram highlights the key points.

*Figure 23: Microservice database - information view*
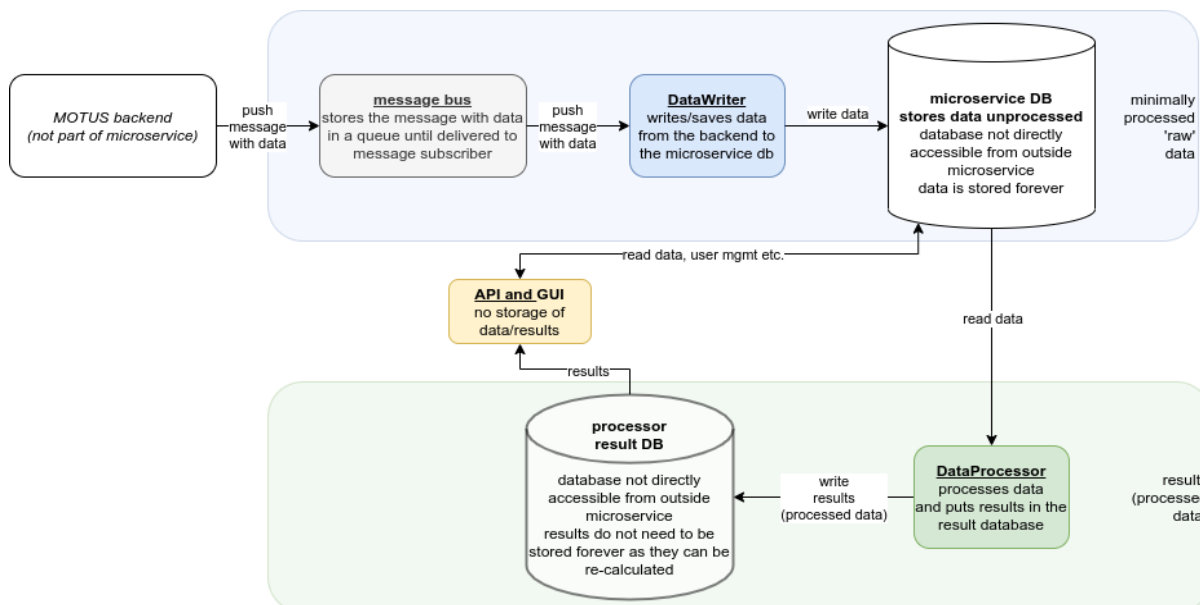
Concurrency view

This view identifies the parts of the microservice that can execute concurrently and how this is coordinated and controlled.

All functional runtime elements are allowed to run in parallel since their responsibilities are clear and non-conflicting (e.g. the DataWriter writes data while the DataProcessor processes data).

The most important concurrency aspect in the microservice architecture are the different DataProcessors which can process data in parallel:

*Figure 24: Microservice database - concurrency view*
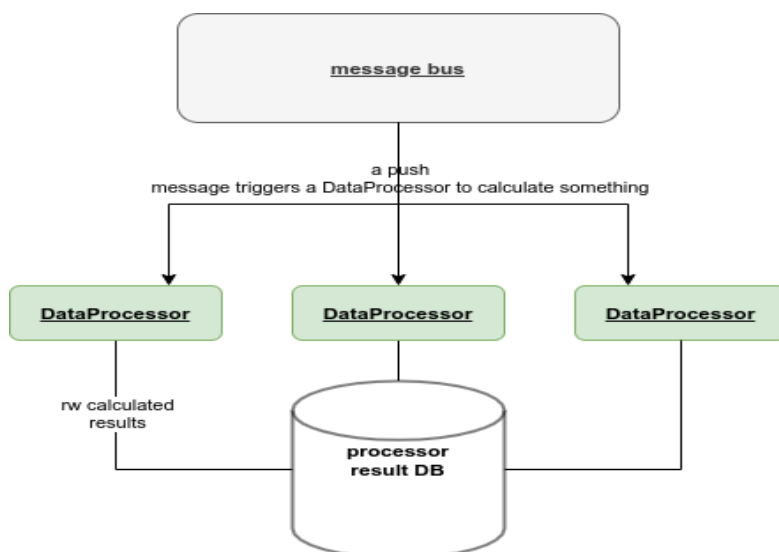
Care must be taken to:

- avoid race conditions: if 2 DataProcessor calculate the same thing, then one DataProcessor might overwrite the results of the other, also if the other's results were more recent.
- avoid unnecessary recalculations: DataProcessors should check the database to make sure the results for its calculation are not already there. In that sense, the processor result DB also acts are a kind of cache. If a single result is a combination of multiple small results, then cache optimizations might be possible by re-using the finer-grained results. E.g. suppose you need to calculate a timelog of a day. If a day is in progress then possibly only recalculating the last hours is enough instead of recalculating the whole day.

Deployment view

A microservice is deployed as a collection of Docker containers: each functional runtime element is built as a Docker container. This makes all elements (almost) independent from the host OS.

Depending on the performance and scalability requirements (see perspective), different deployment strategies are possible. Here are some examples:

Figure 25: Microservice database - deployment view
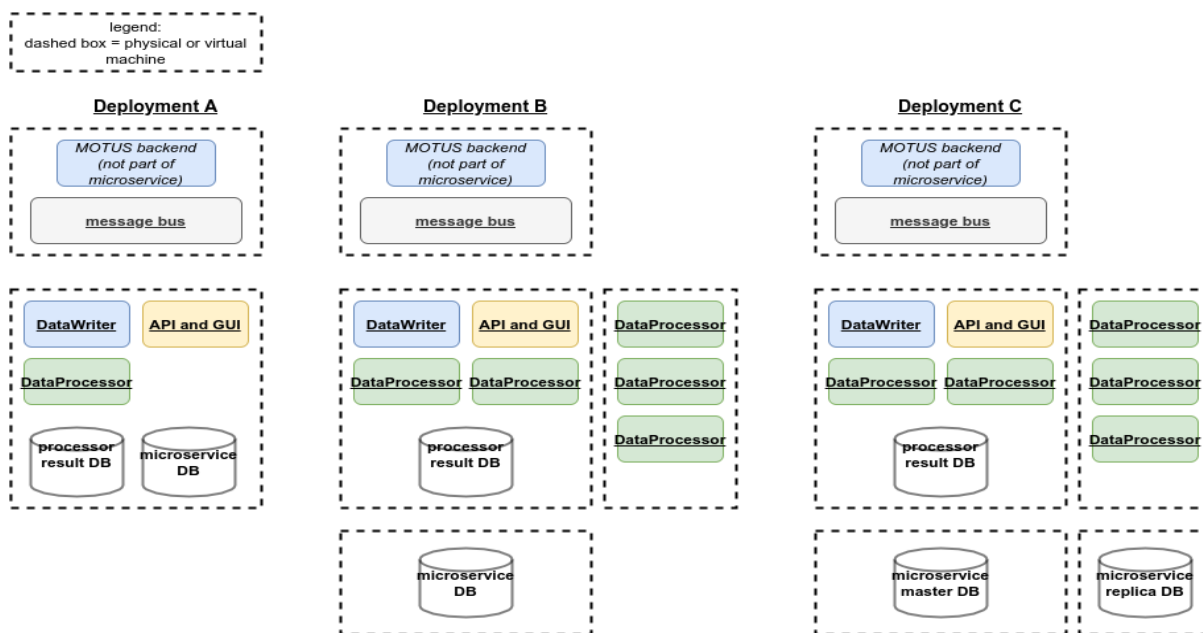
Note that the message bus should always be ready to accept messages from the motus backend (to avoid the backend to be blocked). To exclude networking issues, backend and message bus are on the same machine.

Operational view

Installation and upgrade

The microservice is a collection of Docker containers that will be managed, scaled and deployed with the help of Kubernetes (https://kubernetes.io/).
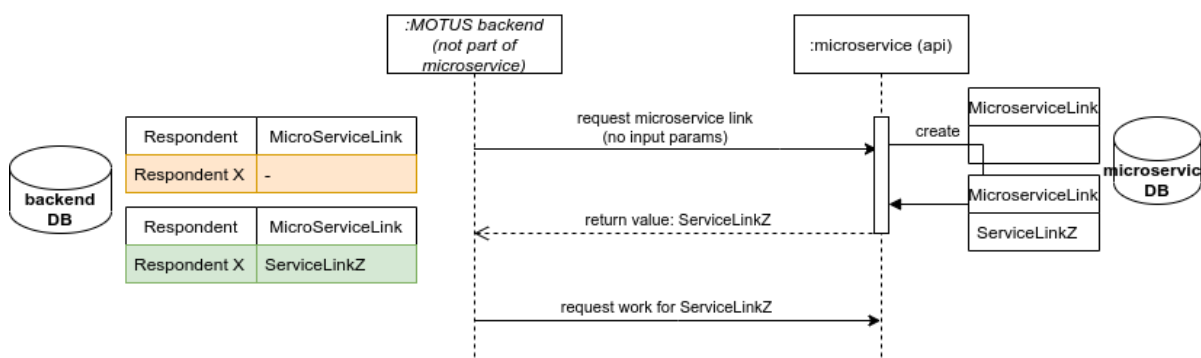
8.4 Perspectives

8.4.1 Regulation perspective

Privacy

Sensitive information must be restricted to the database of the main application. The microservice is not allowed to pull user/respondent private information into its own databases.

The following mechanism is foreseen:

Figure 26: Privacy regulation perspective

The Respondent X entry is never visible in the microservice. Backend and microservice are linked to each other via a "microservice link". The microservice only has knowledge of the abstract "microservice link", which essentially is only an id (uuid, guid…).

If the microservice database would be shared for researchers (e.g. for postprocessing), then the sensitive information of the respondent cannot be leaked.

8.4.2 Performance and scalability perspective

Depending on the application (e.g. type of research) and the type of microservice (e.g. processing intensive vs IO intensive), performance and scaling of the microservice can be tuned as follows:

- DataProcessors can be scaled:
    - by creating multiple instances,
    - by distributing instances over multiple machines,
    - see 'concurrency view' of how they distribute work.
- microservice DB can be scaled:
    - by db replication,
    - by distributing the replication databases over multiple machines,
    - microservice results DB could be scaled similar to the main microservice DB. Since this is probably not a heavily loaded database (no complicated queries, only results), scaling might not be needed.

8.5 GeoService microservice

8.5.1 Architecture

The GeoService architecture follows the generic architecture. The naming of some components is sometimes different though:

- Api and Gui: is a webserver application called 'GeoGui' and is the microservice graphical frontend and api.
- DataProcessor is called GeoProcessor. The Geoprocessor calculates the clusters of when a respondent is in transport / not in transport, and finds the nearby places. It can be started with the command: php artisan geoprocessor:start
- DataWriter is called GeoWriter. The GeoWriter stores geotracking points in the database without any logic. It can be started with the command: php artisan geowriter:start

8.6 Rabbitmq API

8.6.1 GeoWriter

Consumer

Rabbitmq service: geoservice

Topic: trackingpoint.add

Queue: trackingpoint.add

Consumer tag: geowriter

8.6.2 GeoProcessor

Consumer

Rabbitmq service: geoservice

Topic: tracker.process

Queue: tracker.process

Consumer tag: geoprocessor

Body: {"tracker_uuid": <tracker uuid>}

Producer

Rabbitmq service: geoservice

Topic: tracker.processed

Body: {"tracker_uuid": <tracker uuid>}

8.7 API

The API documentation of the GeoService backend is part of the code repository. (storage/api-docs.yaml).

This work work package relies on previous work done under contract for Eurostat and the project 'Innovative Data Collection Solutions for HBS and TUS (Specific Contract Number: 000057 Framework Contract Number: 07112.2016.003-2017.018; Lot 1).

Within this project further modifications are done to the plugin and are reflected in the submission of the 'geotracker-croess-D3.2-M4' package.

If the microservice is running, the documentation can be browsed and the endpoints can be tested from within the microservice itself. Only administrators are allowed to view this documentation so they should be logged in to the backend first.

If running on a localhost the documentation can be found here: http://127.0.0.1:8000/api/documentation

Figure 27: API documentation of the GeoService

8.8 ReceiptScanner microservice

8.8.1 Architecture

The ReceiptScanner architecture follows the generic architecture with the following differences:

- Api and Gui: the receiptscanner has no a GUI, only an API.
- DataProcessor is called ReceiptProcessor. It can be started with the command: php artisan receiptprocessor:start
- DataWriter: not required for the receiptscanner because there is no continuous flow of respondent data (in contrast to the geotracker which continuously receives geotracking points). The DataProcessor receives the receipt image file location via a rabbitmq message.

8.9 Rabbitmq API

8.9.1 ReceiptProcessor

Consumer

Rabbitmq service: receiptscannerservice

Topic: receipt.process

Queue: receipt.process

Consumer tag: receiptprocessor

Body: {"scanner_uuid":<scanner uuid>, "receipt_uuid": <receipt uuid>, "timestamp": <timestamp>, "receipt_filename", <filename>}

Producer

Rabbitmq service: receiptscannerservice

Topic: receipt.processed

Body: {"scanner_uuid":<scanner uuid>, "receipt_uuid": <receipt uuid>}
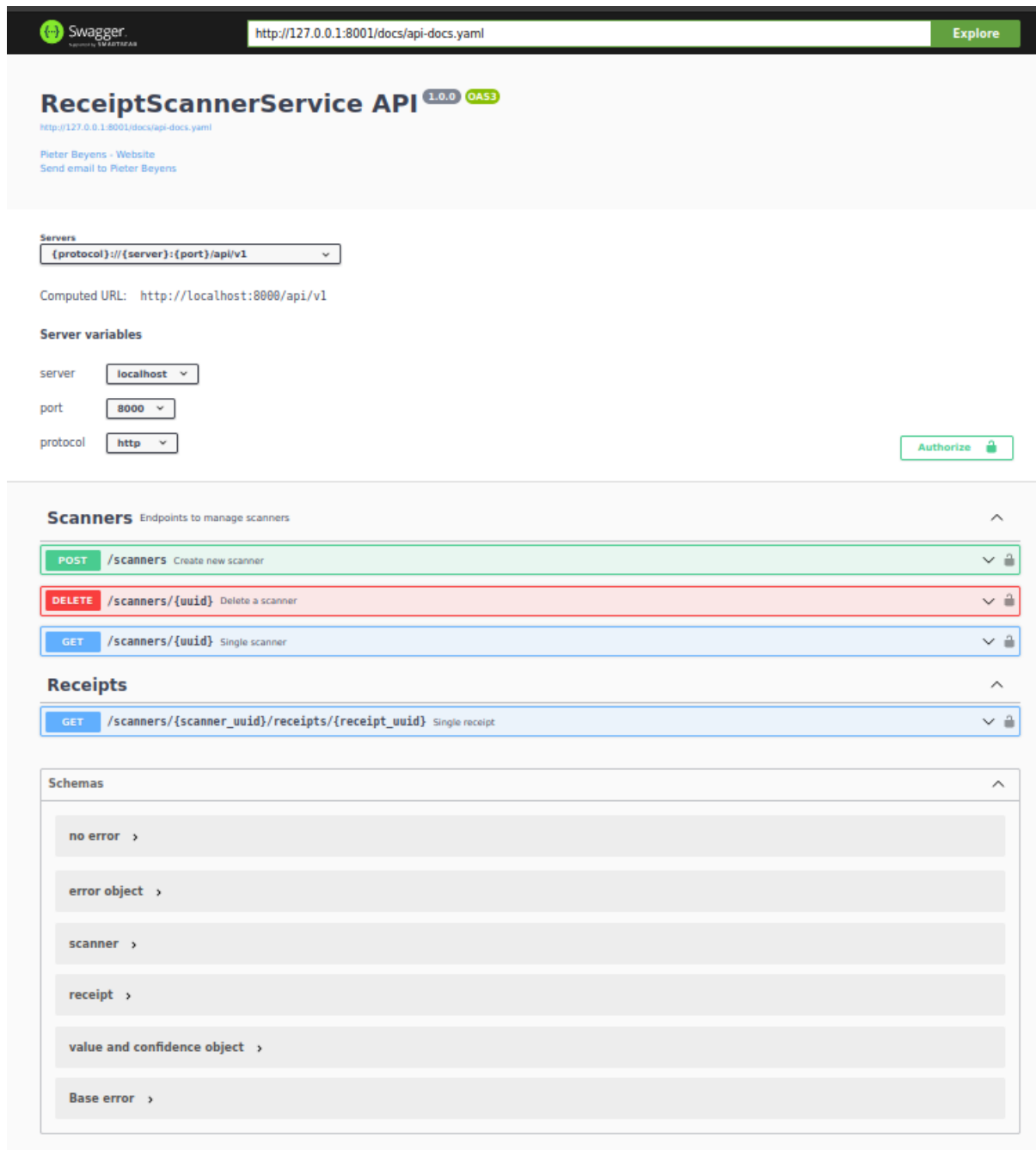
8.10 API

The API documentation of the receiptscanner is part of the code repository. (storage/api-docs.yaml).

The package has been delivered as the 'receiptscanner-croess-D3.3' package.

If the microservice is running, the documentation can be browsed and the endpoints can be tested from within the microservice itself. Only administrators are allowed to view this documentation so they should be logged in to the backend first.

Figure 28: API documentation of the Receipt Scanner

If running on a localhost the documentation can be found here: http://127.0.0.1:<port>/api/documentation

8.11 R

8.11.1   Introduction

The R setup is different from the architecture for geoservice and receiptscanner. CPU-intensive database tasks such as R analysis are offloaded to a server which is dedicated to analysis. As such, the main database is not overloaded by complicated analysis queries.

This architecture reflects the well-known difference between 2 types of data processing systems:

- OLAP or online analytical processing, vs
- OLTP or online transaction processing.

The motus server is an OLTP system : simple and fast queries to the database, while the analysis server is an OLAP system: slow, complicated queries to the (replication) database.

The motus server will request an R analysis and will fetch the results when available.

Figure 29: R - MOTUS server
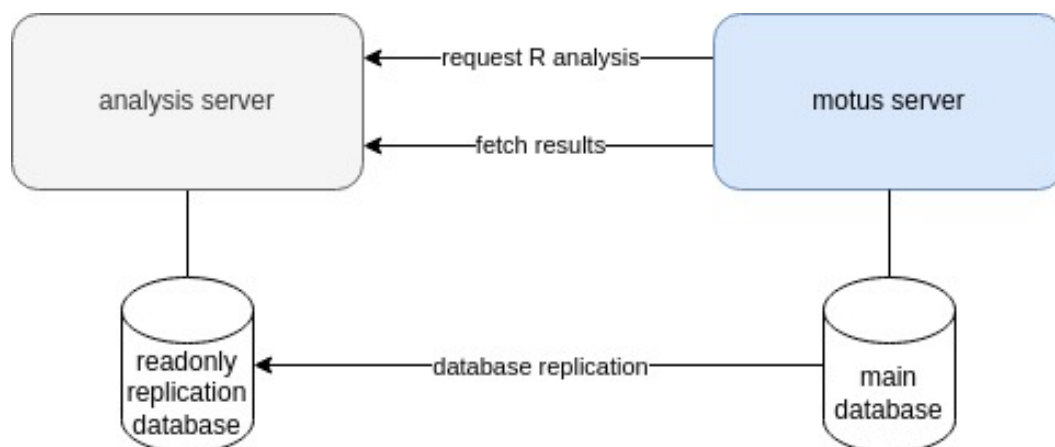
The package has been delivered as the 'libmicroserviceR-croess-D3.4' package.

9 WP3 – D3.2 – Geolocation microservice

Overall goal: live demo, documentation and source code of geolocation microservice in GitHub.

9.1 Introduction

This work work package relies on previous work done under contract for Eurostat and the project 'Innovative Data Collection Solutions for HBS and TUS (Specific Contract Number: 000057 Framework Contract Number: 07112.2016.003-2017.018; Lot 1).

Within this project further modifications are done to the plugin and are reflected in the submission of the 'geotracker-croess-D3.2-M4' package.

In particular changes have been made to the GeoSerice package:

- Integration of application code in a microservice. The microservice consists of 3 Docker containers which are responsible for writing the tracking points to database, processing the tracking points and accessing the tracking points from the MOTUS core.
- Integration of the microservice in a core platform. This includes support for asynchronous (RabbitMQ) and synchronous communication (REST json).
- Support for deployment of the microservice in a kubernetes cluster with helm.
- Changes to support the microservice privacy concept..
- Improved clustering algorithm which determines whether a user is in transport or not.
- Mapping of places (as found with Google Places) to more generic openstreetmap categorization. From this categorization, TUS activities could be derived in the core platform.

This document describes the microservice called GeoService. First the document explains its Backend i.e. how a user (admin) accesses and manages the microservice. This document is a copy of the earlier submitted project delivery, but for means of easier reading added here as well. The second part add GeoService specific information to the general information on the microservice architecture as being described in Deliverable 3.1 (to which also the API documentation can be found).

9.2 GeoService Backend

The backend of the geolocation microservice *GeoService* has multiple parts, which are being outlined below.

9.3 Login

Once the microservice is installed a user with administrator rights can be added to the database. This user is able to sign into the geolocation microservice using an email address and a corresponding password.

Additional users can be added by the admin user.
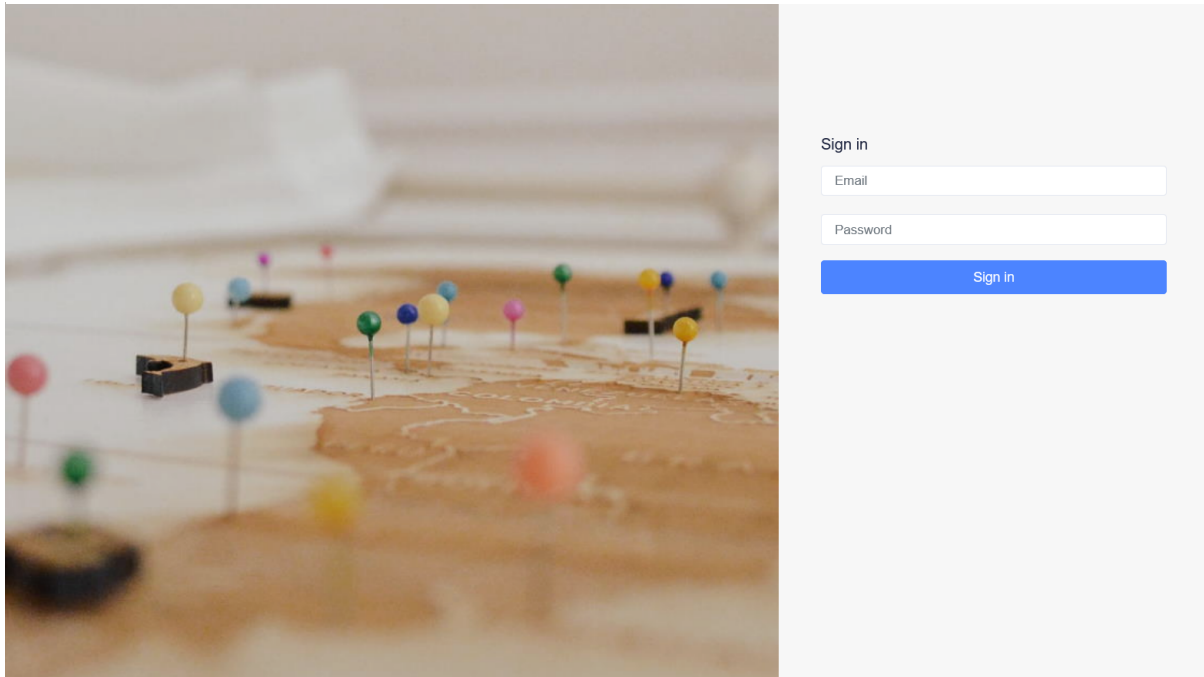
The number of users that can be added is unlimited. All have admin rights. Access to the geolocation microservice should only be granted to persons who want to check and/or validate data collected via the microservice.

9.4 Dashboard

The dashboard gives a simple overview of the collected data.

The left side sidebar can be used to access the different parts of the backend:

- Home: Shows the dashboard
- Trackers: All devices that allow tracking
- Locations: All public locations
- Users: User management

Figure 31: Microservice dashboard
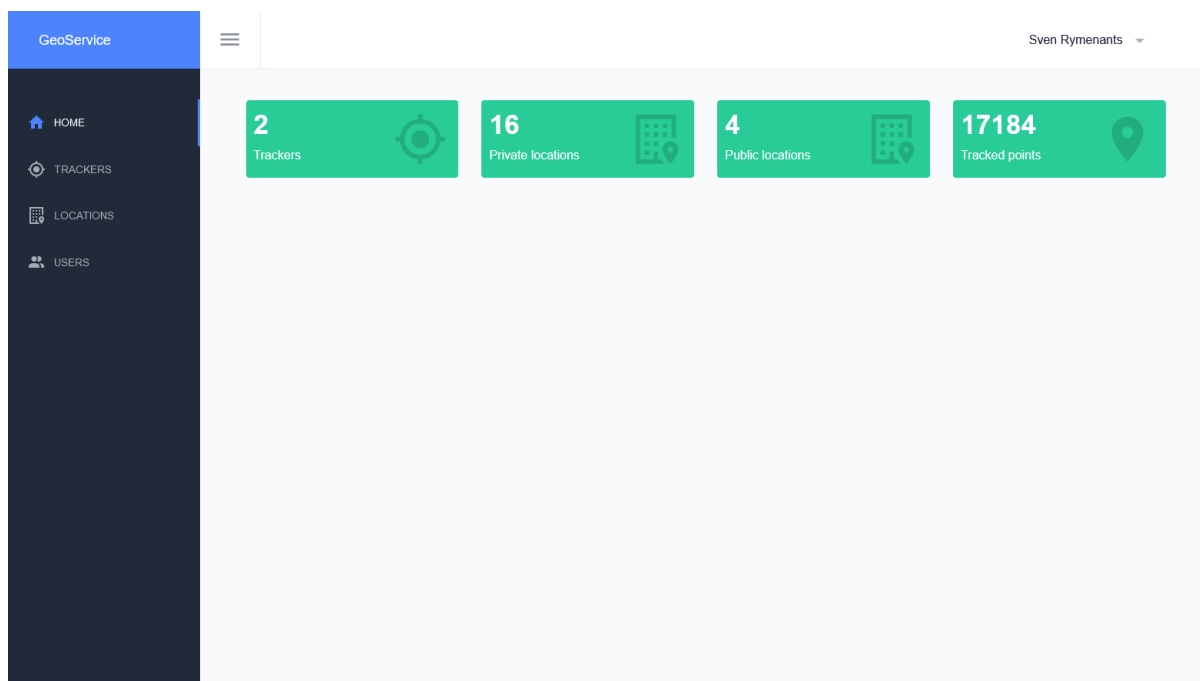
9.5 Trackers

Trackers are identifiers for the different devices that have the mobile geolocation (Cordova) plugin installed.

The given UUID is an anonymous reference to a respondent. All tracked points and private locations of the respondent can be accessed from this screen.

Figure 32: Microservice dashboard trackers
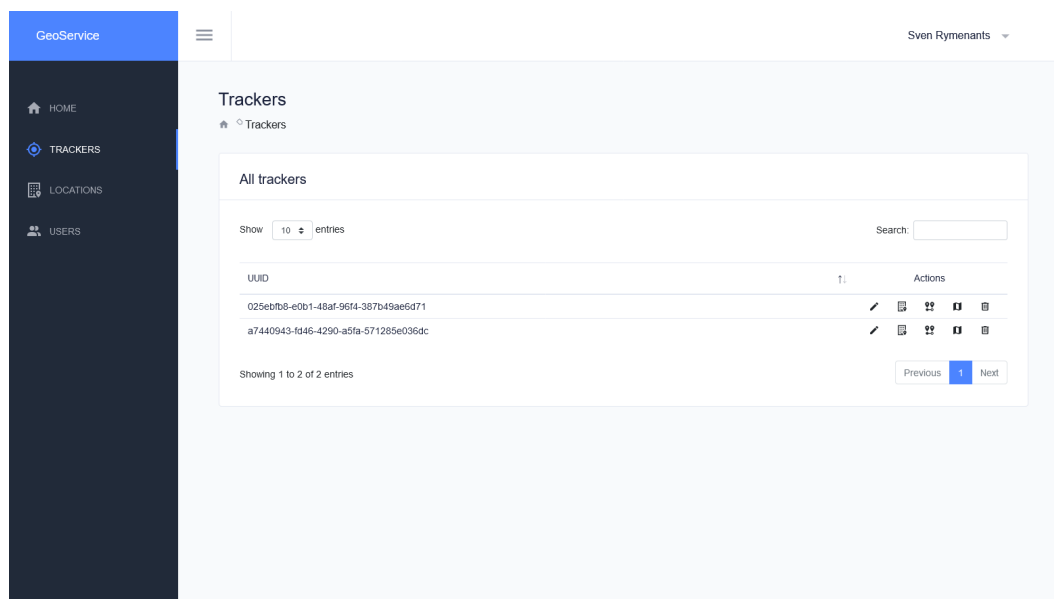
When a respondent is created in the application layer this should make a call to the microservice to register the respondent. As a result a UUID of the new tracker will be returned to the application which should be used to access the services offered by the GeoService microservice.

The icons in the list give access to more details of the tracker:

- Pencil icon: edit tracker information
- Building icon: edit tracker private locations
- Two markers icon: view a list with all tracked points
- Map icon: show all tracked points on a map
- Trash icon: delete the tracker (including logged data)

In the list one can view the UUID for the tracker, since the backend of the microservice offers a view on the database, the UUID is the only property/column that can be edited in the edit view (pencil icon).

Figure 33: Microservice dashboard tracker editing
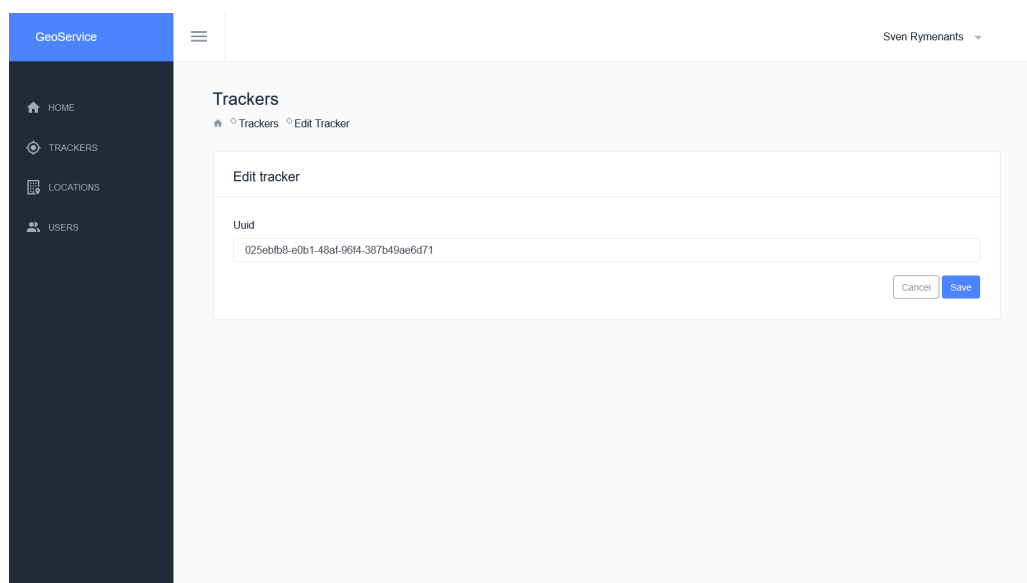
The building icon in the list shows a list of all tracker private locations. This feature is available through the API of the microservice, so it can be used by a mobile app, like the MOTUS-app is an example.

Figure 34: Microservice dashboard tracker building
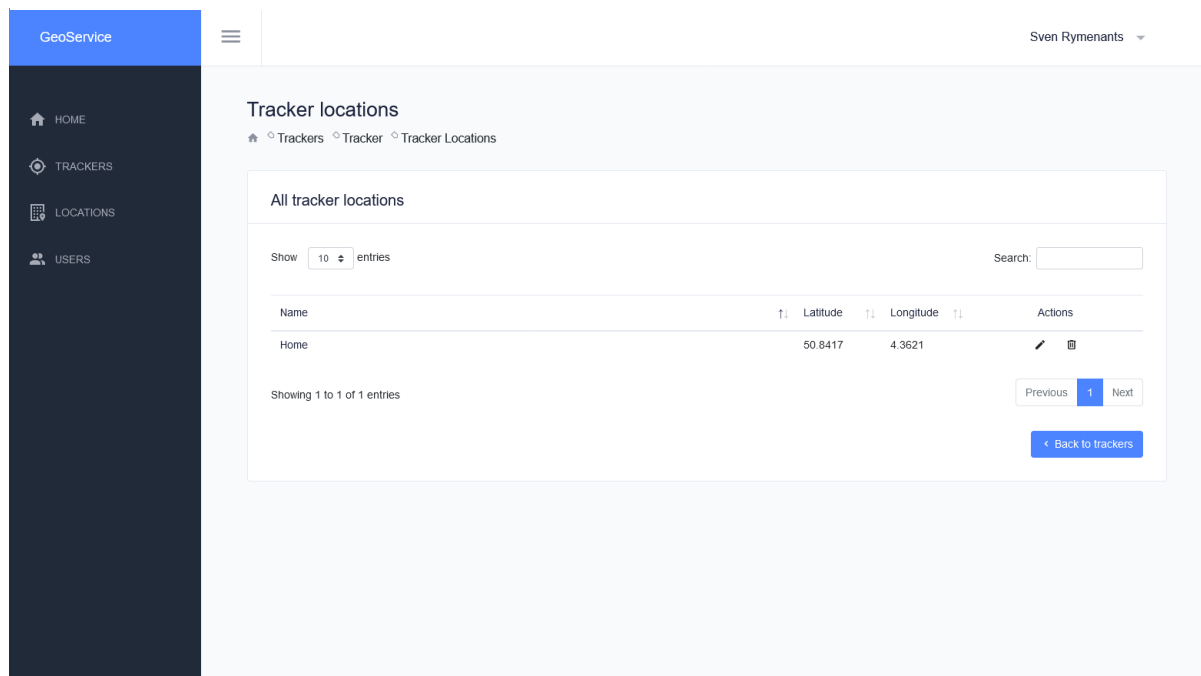
Adding, editing or removing locations can be done in the backend, but also using the API of the microservice.

An app that uses this microservice can use this feature to add location management for the respondents without the need of implementing a full blown location management system.

Using this feature adds another advantage, as the near feature of the microservice will include the private locations of respondents next to the existing lists that are retrieved from 3th party services like Google Places, OpenstreetMap Overpass or Foursquare Places. The communication with these services is only done based on the UUID, so privacy of the tracker is protected.

When editing a point in the backend it is possible to add or change the name of the location. To change the middle or radius of the geofence, the marker can be dragged in the interactive map.

Figure 35: Microservice dashboard tracker location management
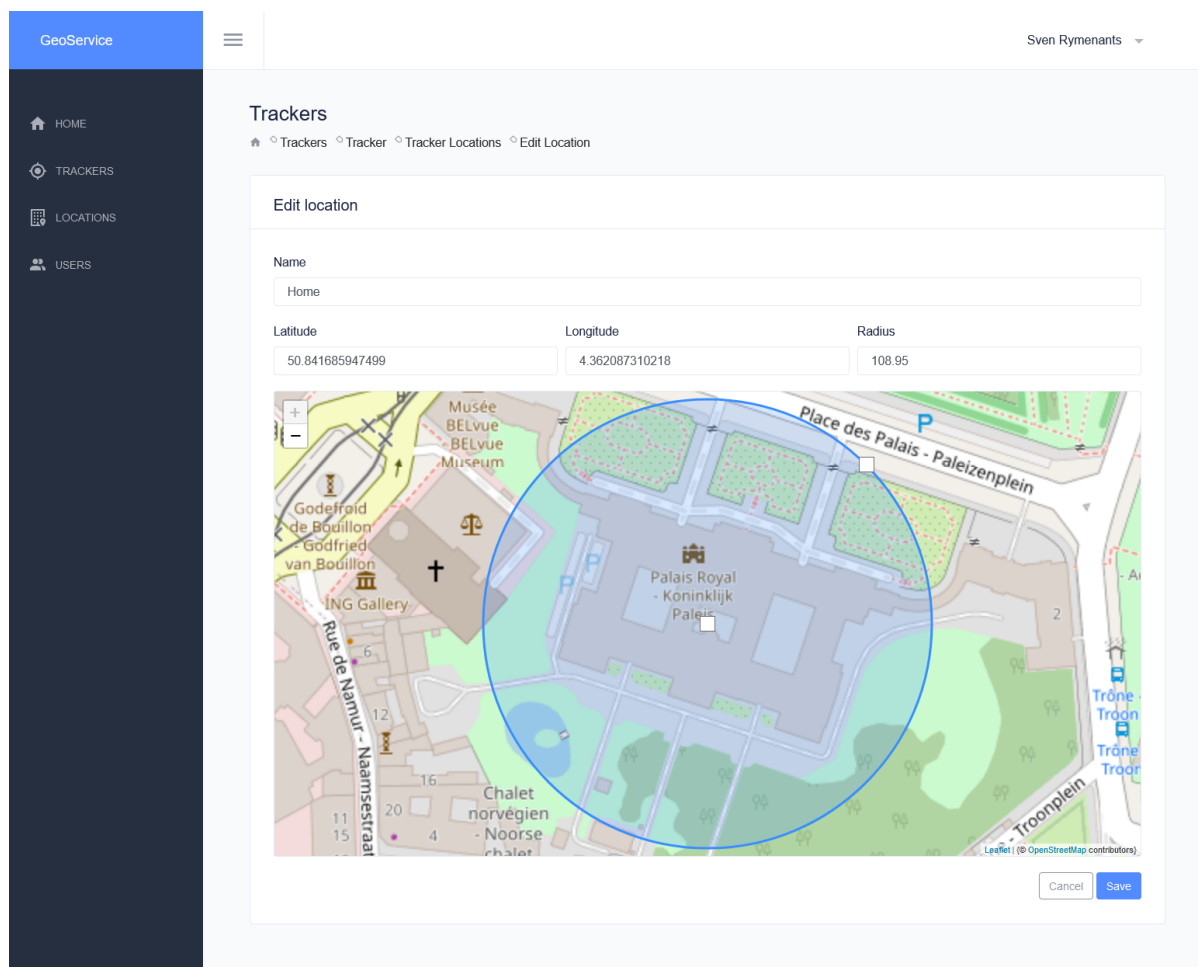
The tracking points can also be viewed as a list. All logged tracking points for the selected tracker are visible and can be browsed from this list. The show map the points on a map the user can click on 'Show map'.

For each point, there are 2 possible actions:

- Eye: view the point on a map
- Trash: remove the point from the list

Figure 36: Microservice dashboard tracking point
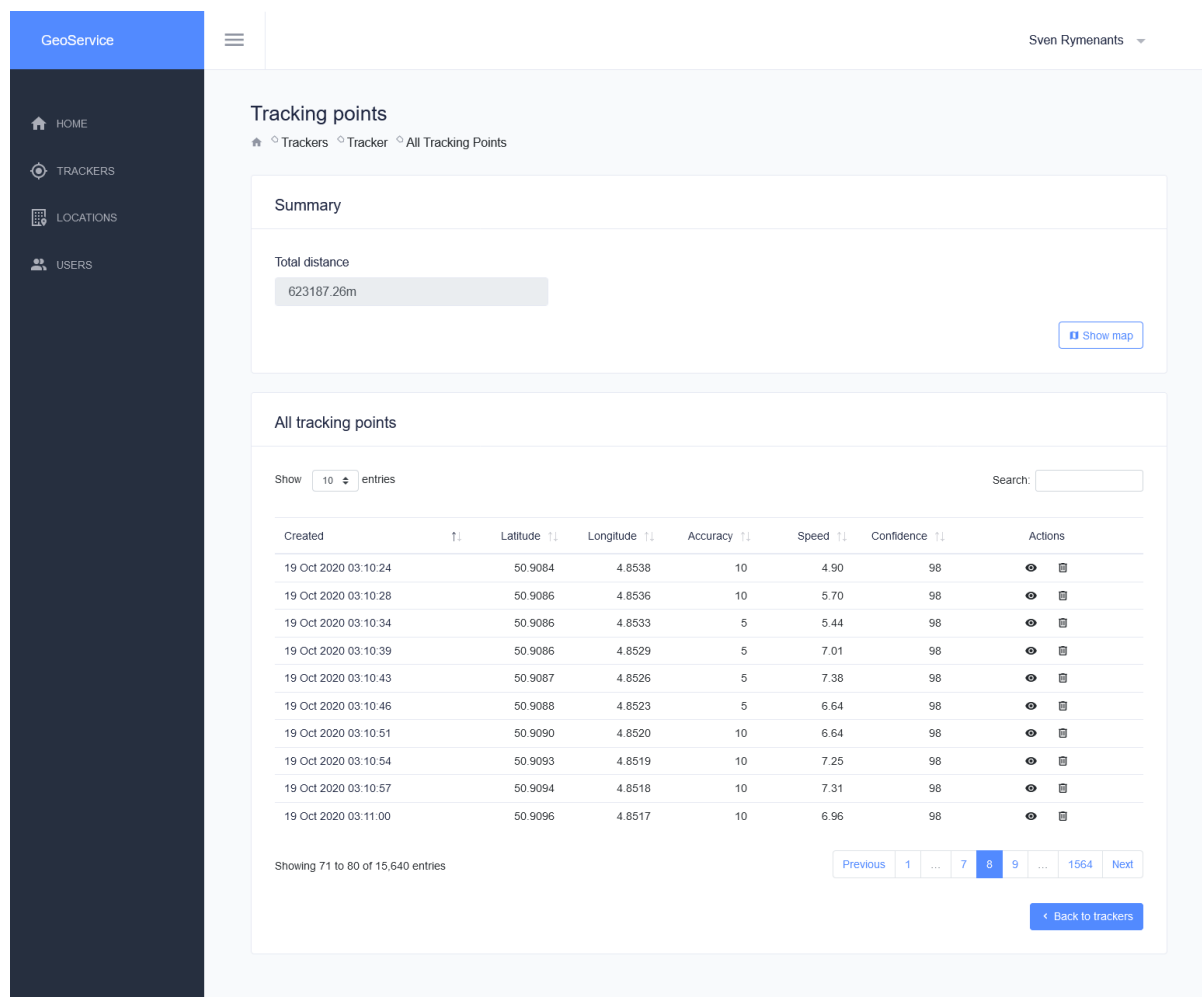
The map view shows the path of the tracker. Each marker on the map is clickable and allows the creation of a new location.

The visible path contains only the points that are logged during the selected day. It is possible to select another day using the date selector.

To increase usability, the picker highlights dates when logged data is available.

Figure 37:Microservice dashboard logged data
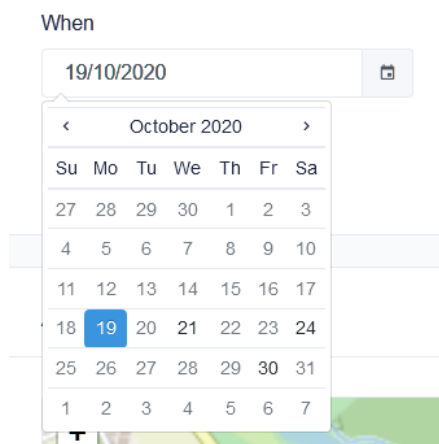
The near button will show a list of all known nearby locations of the selected points using one of the enabled services.

Figure 38: Microservice dashboard tracker - near button
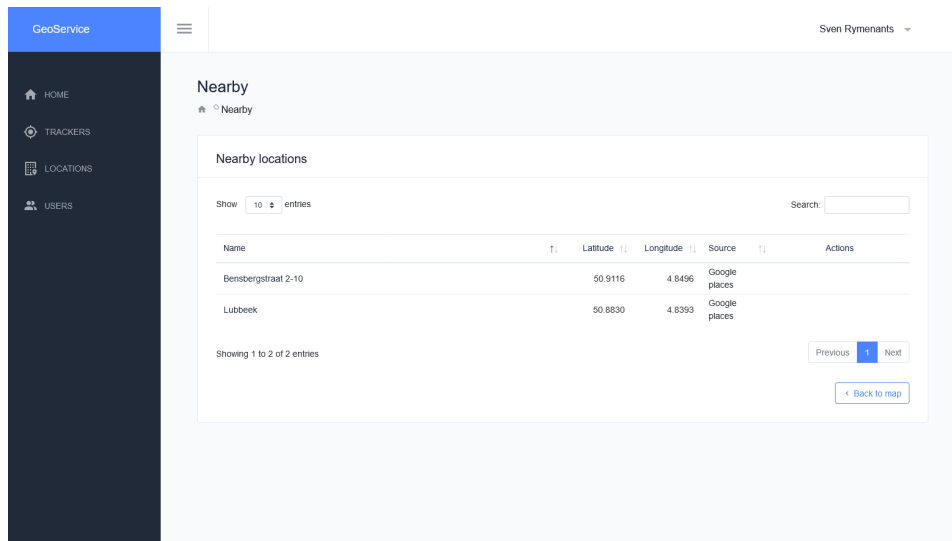
The arrows on the path indicate in which direction the tracker moved from one point to another.

Figure 39: Microservice dashboard direction indication
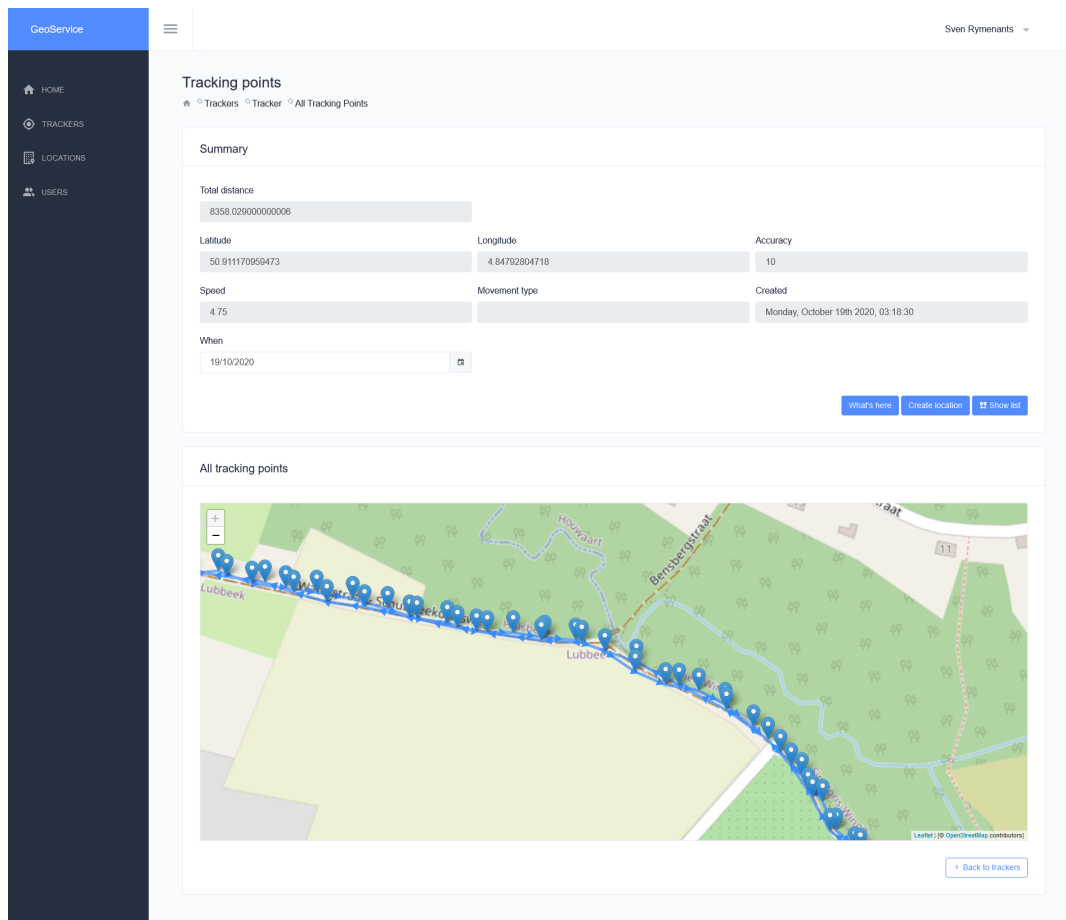
The what's here button shows a list of all places that can be found near the selected point.

In the list it is possible to view the data source of the found nearby locations.

Note that there are no actions possible for the listed locations, this is due legal limitations of the API of Google and/or Foursquare as they disallow the caching or storing of retrieved places in another system that uses their API.

Also note that for each request a small fee might be charged by the suppliers. This can be avoided when using OpenstreetMap Overpass, but it has its limitations on the input side.

Another option is to click the create location button which will create a new (private) location for the tracker.

When creating a new location, entering a name is required, whereas latitude, longitude and the geofence radius can be altered before creating the location.

Figure 40: Microservice dashboard creating new location
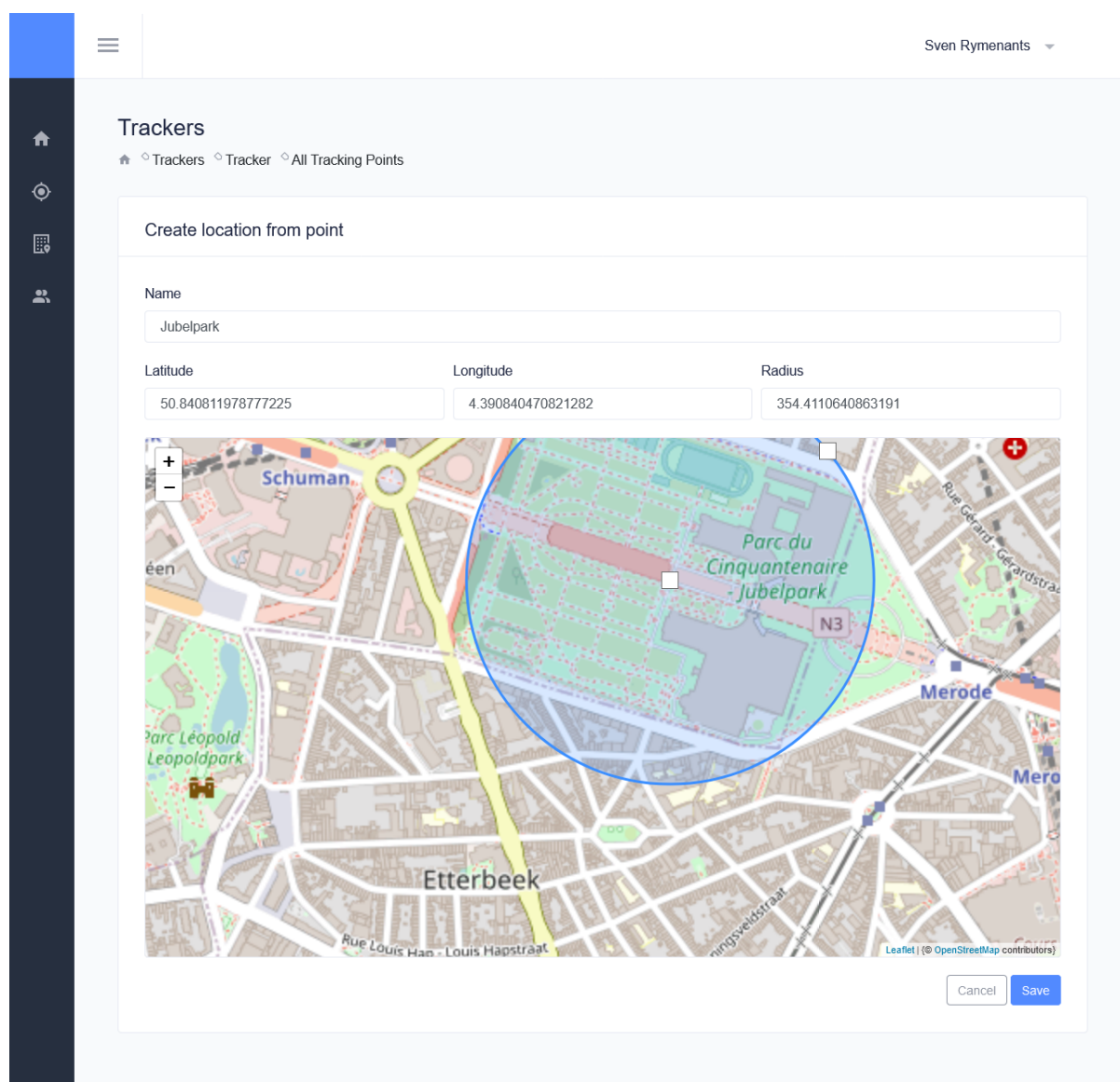
After saving the location, this location will be only accessible for the tracker. The same functionality can be made available in a mobile app.

9.6 Locations

If required it is also possible to create locations that are available to all trackers. This can be done in the Locations menu.

The usage can be compared to the views in the private locations views. There is a list with locations and each location can be altered or removed using the pen or trash icon.

Figure 41: Microservice dashboard locations menu
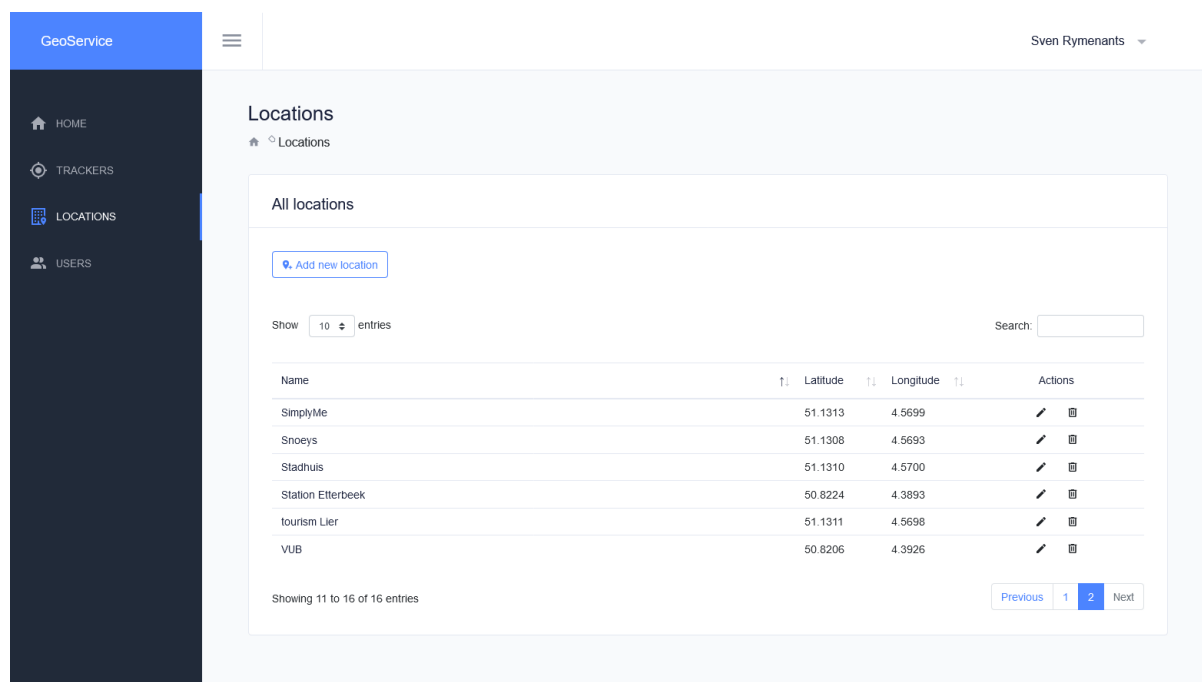
The locations in the list will also be made available to the nearby endpoint in the API, at the moment a near request is issued.

9.7 Users

The users are the administrators of the microservice GeoService. These users are allowed to log in to the backend and to manage all captured content.

The list has two actions:

- Pencil icon: edit user
- Trash icon: delete user

Figure 42: Microservice dashboard users menu
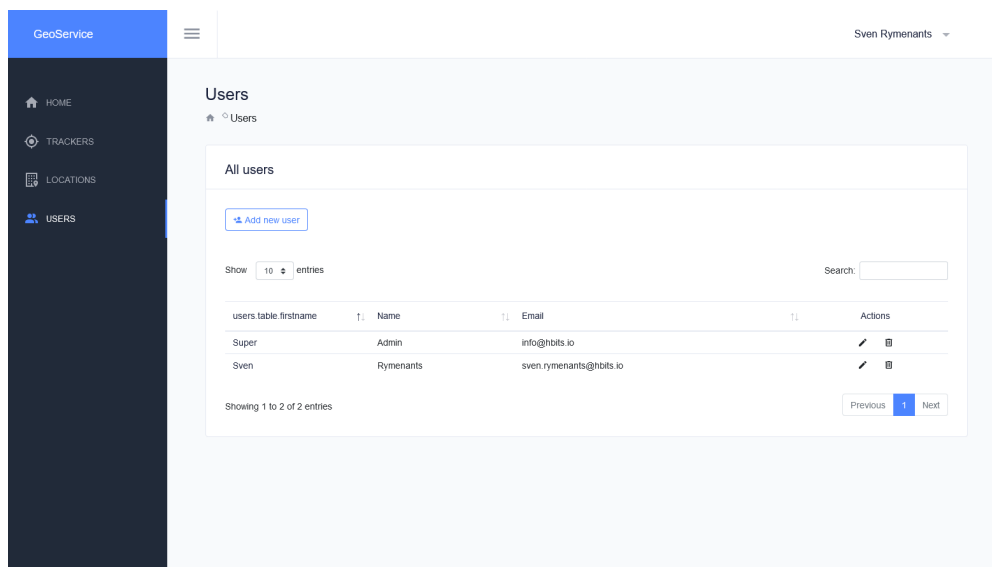
Note: The users cannot be managed through the API.

When editing a user the first and last name can be altered.

The access to the backend can be accomplished with an email address and a password. On this page each user can be managed.

Figure 43: Microservice dashboard management of users
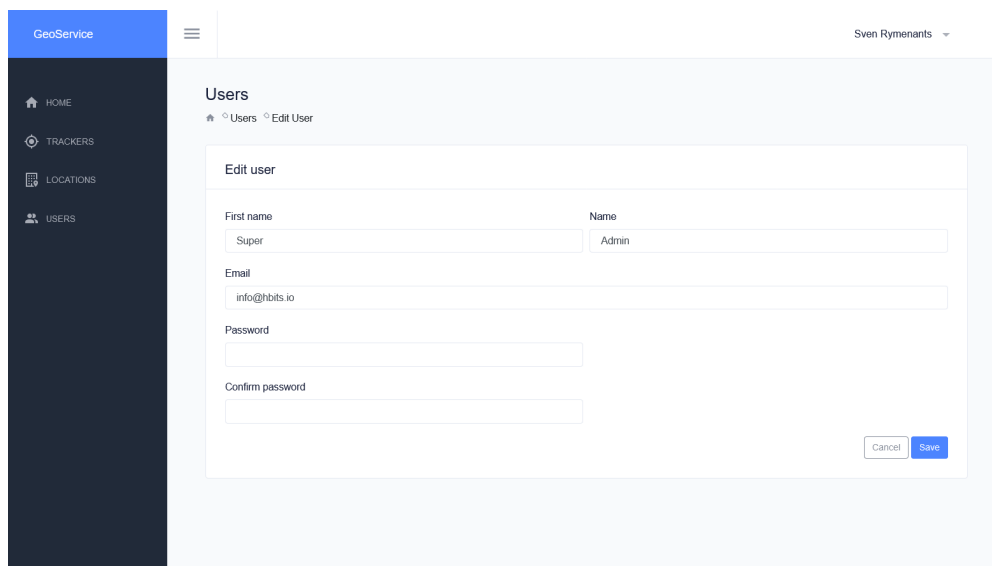
9.8 Profile

The last view is the profile. On the right side of the header there is a drop down menu when an administrator clicks his/her name. From there the profile page can be reached.

Figure 44: Microservice dashboard profile
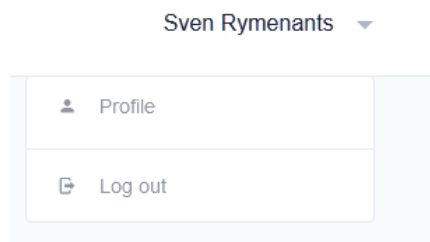
In the first tab administrators can change name, email and password.

The second tab contains information about the active and past sessions. The tab contains information about the location and device used for all sessions.

Figure 45: Microservice dashboard profile information
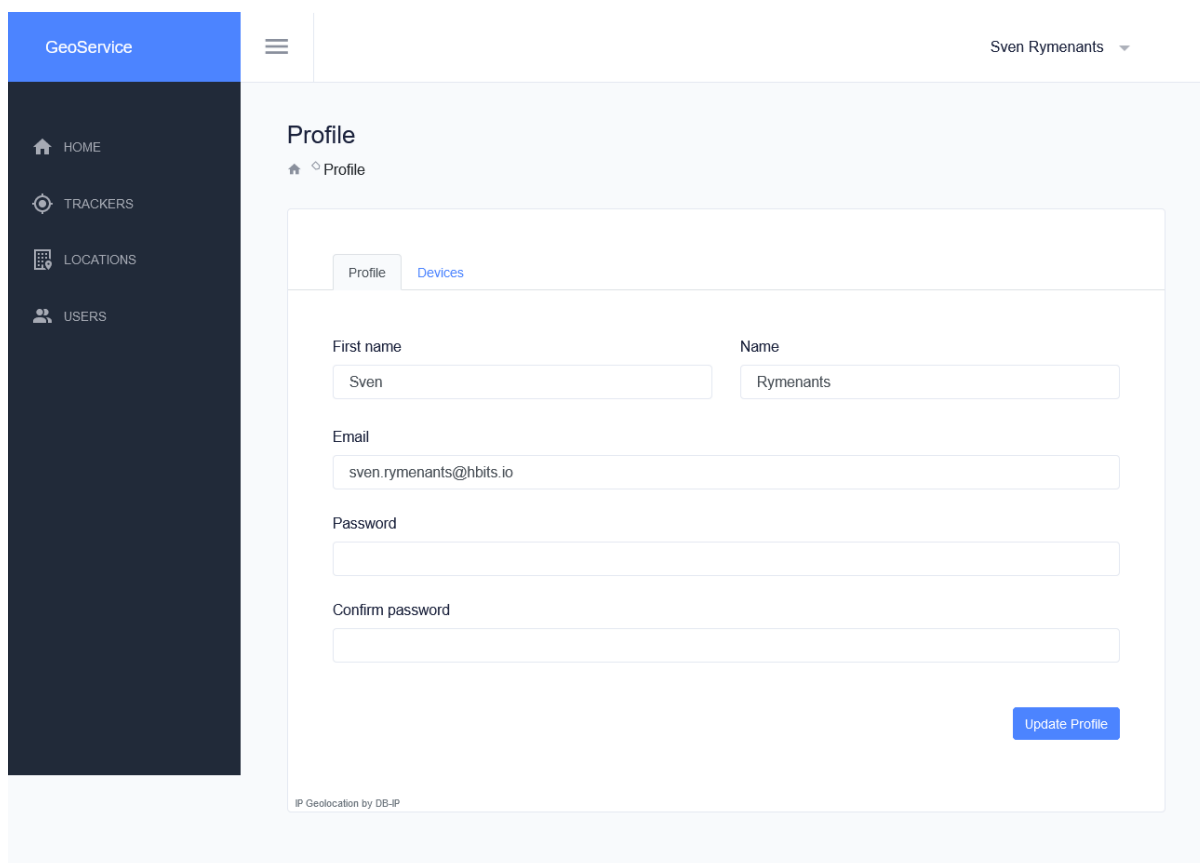
Older sessions can be removed using the Log out from device button.
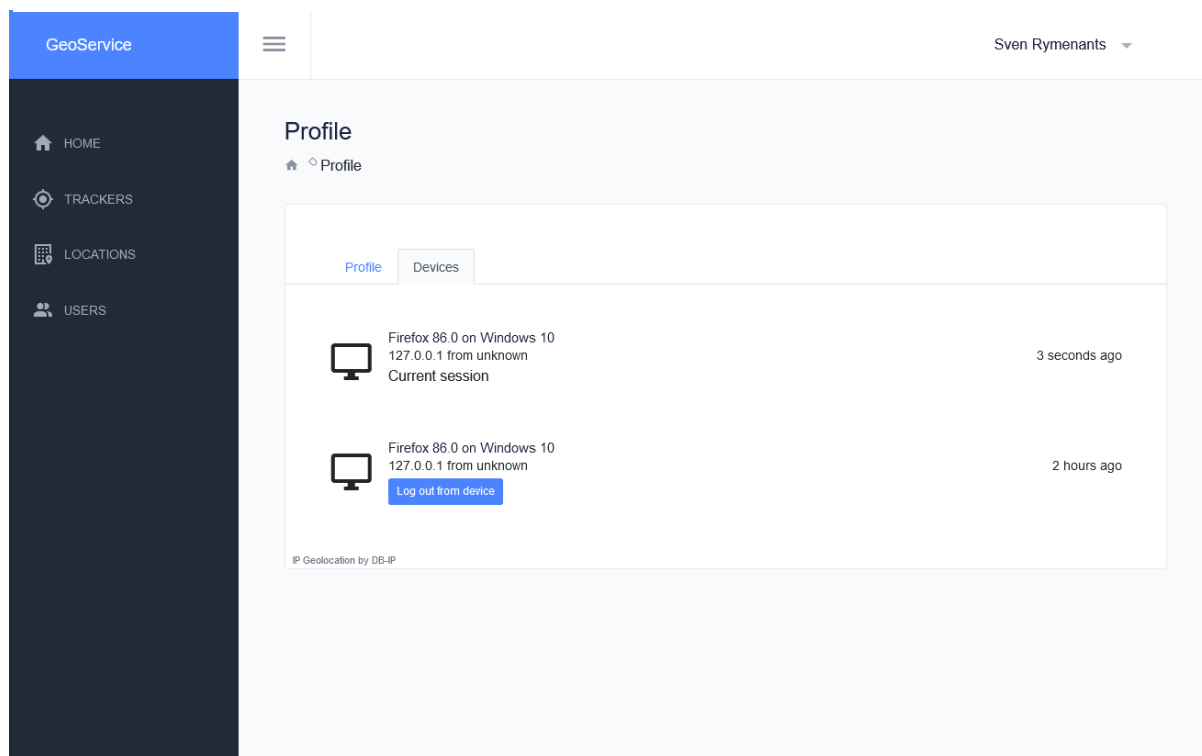
9.9 GeoService architecture

The GeoService architecture relies on the general microservice architecture as being handled in Deliverable 3.1 describing the software architecture, database and API for the 3 developed microservices.

Below we give further detail on the contents and figures reported in Deliverable 3.1, and structured via the different views.

Context view

As shown in Figure 21 the GeoService fits between 2 elements, being the user and the system. Although in most diagrams the MOTUS backend is used to illustrate interaction with the microservice, any other application could use the microservice for data processing tasks. The microservice is independent and not strongly coupled to the MOTUS backend.

There are 2 types of users: the respondent and the researcher.

Respondents who participate via the MOTUS front office (or via another tool). Respondents don't experience the in between position of the microservice.

Researcher interact with the service via the back-office of the GeoService (see above), and with the back-office of MOTUS or a different adopting environment of the microservice.

As for all microservices developed within the context of this project there is no direct link between the microservices and the backend db. In the case of MOTUS it means that the MOTUS backend has full control over which data is delivered to the microservices.

On the functional level Deliverable 3.1 described the different functional elements of the microservice architecture in general: Message bus, API and GUI, Data Writer, DataProcessor, Processor Result DB, Microservice DB (see Figure 22). With the GeoService

some functional elements (and responsibilities) of the overall microservice are extended. These are being discussed below.

API and GUI

The GeoService also needs information (e.g. respondent's private locations) which is stored in the MOTUS backend. This is handled as illustrated in the diagram:

Figure 47: GeoService architecture - handling in MOTUS backend
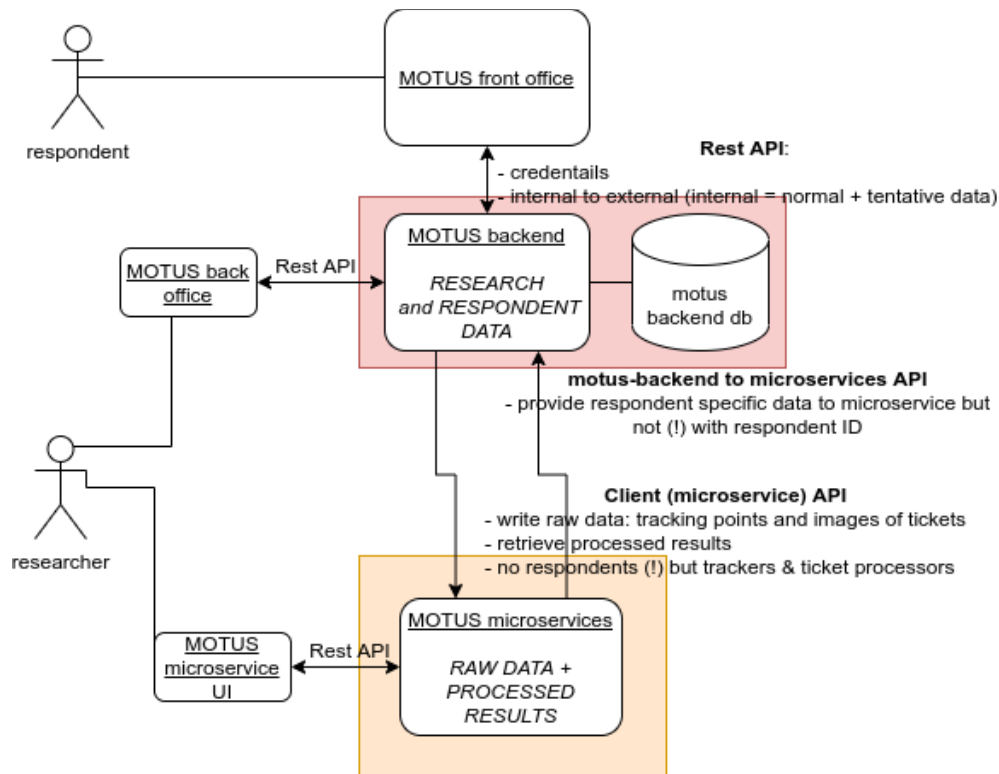
The MOTUS backend exposes an API which can be used by the GeoService microservice to retrieve tracker information. The MOTUS backend is under full control of the respondent's data is only provides non-sensitive data to the microservice.

DataWriter

In the GeoService, this component adds geotracking points to the microservice database, without any intelligent processing.

DataProcessors

Processing geotracking points is done in the following steps:

- clustering: divide geotracking points into clusters, define which clusters are in transport (e.g. car, train…) or not (staying for some time at a certain location)
- for the clusters that are not in transport, request nearby locations: sources of locations can be Google Places (Places API), OpenStreetMaps (Overpass API), FourSquare (Places API) but the private locations of the respondents as well.
- for all found location/places, map them into well-defined categories. The MOTUS backend is then able to map those location categories to activities which is the main interest for TUS.

Information view

From an information point of view the GeoService, just like the other microservices, stores, manipulates, manages and distributes information. The diagram highlights the key points (see Figure 23).

For the GeoService database however the tracker table holds all trackers that are registered in the microservice.

The table with tracking points contains a list with all tracked points linked to a tracker.

Furthermore all locations that are private are also linked to the tracker itself.

Public locations are stored in the same table, but will not be linked to a tracker.

All users are stored in their own table, any sessions linked to the user are stored in a separate table.

Since the microservice uses a relational database to store all data, tracking points and locations will be removed if a tracker is removed from the database.

Figure 48: GeoService database structure
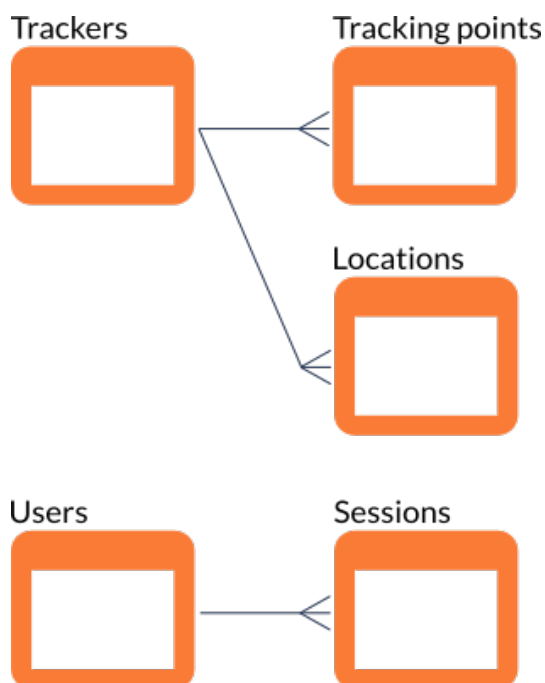
Database tables

Tracker table

The ID of the tracker is for internal use, for all interfacing to in the backend and API the UUID is used.

Table 7: GeoService tracker table

id	ID of the tracker
uuid	UUID of the tracker

Trackingpoint table

The ID of the tracking point table is for internal use, for all interfacing to in the back end and API the UUID is used.

The tracker_ID links to the tracker in the tracker table.

Table 8: GeoService trackpoint table

id	ID of the tracking point
uuid	UUID of the tracking point
tracker_id	Link to the tracking point
created_at	Creation date and time of the tracking point
latitude	Latitude
longitude	Longitude
accuracy	Accuracy when the point was created
speed	Current speed when the point was created
confidence	Confidence in the detected activity type
activity	Activity type enum value

Locations table

The ID of the locations table is for internal use, for all interfacing to in the backend and API the UUID is used.

The tracker_ID links to the tracker in the tracker table.

Table 9: GeoService locations table

id	ID of the location
uuid	UUID of the location
created_at	Creation date and time of the location
updated_at	Last update date and time of the location
name	Name of the location
longitude	Longitude
latitude	Latitude
radius	Radius of the geofence
visible	Visible in public queries
tracker_id	Link to the tracker, or null for public trackers

User table

The ID of the users table is for internal use, for all interfacing to in the back end and API the UUID is used.

Table 10: GeoService users table

id	ID of the user
uuid	UUID of the user
name	Name of the user
firstname	First name of the user
email	Email of the user
password	Encrypted password
remember_token	Remember token
created_at	Creation date and time of the location
updated_at	Last update date and time of the location

Sessions table

The ID of the users table is for internal use, for all interfacing to in the backend and API the UUID is used.

The user_ID links to the user table.

Table 11: GeoService sessions table

id	ID of the session
user_id	Linked user
ip_address	IP address of the session
user_agent	User agent identifier string
payload	Session information
last_activity	Timestamp of last activity on this session

Concurrency view

This view identifies the parts of the microservice that can execute concurrently and how this is coordinated and controlled. This is equally the same over all microservices, and also for the GeoService.

Deployment view

The GeoService microservice is like the other microservices deployed in a Docker container where each functional runtime element is build as a Docker container giving way to different deployment strategies.

10 WP3 – D3.3 – Receipt scanning microservice prototype

> Overall goal: live demo, documentation and source code of the receipt scanning prototype in GitHub.

10.1 Introduction

The receipt scanning microservice is an important new component to support HBS. Respondents will be able to take a picture from a receipt. That picture is being sent over – via the MOTUS backend - to the receipt scanning microservice. The microservice processes the receipt image by extracting shopping information which can then be linked to COICOP codes in the backend. Finally, all that information is returned to the respondents as tentative data. Tentative data can be edited and needs to be confirmed by the respondents before it is accepted as valid research input.

The microservice performs 3 tasks:

- *pre-OCR processing*: preparation of the receipt image for Optical Character Recognition (OCR)
  - input: receipt image
  - output: black/white image which is cropped to the receipt itself, in which noise has been removed etc.
- *OCR processing*: Optical Character Recognition
  - input: 'clean' receipt image
  - output: raw text
- *post-OCR processing*: finding shopping information in the OCR text
  - input: raw OCR text
  - output: expense item rows, date, store etc.

It can be deployed as a Docker image.

10.2 Architecture and design

For an detailed explanation of the MOTUS microservices architecture, please refer to deliverable *D3.1 microservices for TUS and HBS*. This section will only describe receipt scanner microservice specifics.

10.2.1 Overview diagram

Figure 49: Receipt scanning microservice overview diagram

MOTUS backend
(not part of microservice)

put file

push message

push message

call

sftp server
exposes receipt images

rabbitmq message bus
asynchronous interface between MOTUS backend and microservice

receiptscanner-api: request results, add users...
API provides a **synchronous interface**

push message

push message

receiptscanner-processor
pre-OCR, OCR, post-OCR

get file

ro calculated results

rw data

rw calculated results

ro

redis processor result DB

microservice DB master

10.2.2 Components

Non-Hbits components:

- *Sftp server*: makes receipt images available to the microservice.
- *Rabbitmq message bus*: used for asynchronous communication between backend and microservice. Messages are: requests for processing and notifications when processing is finished.
- *Mariadb microservice database*: stores scanners (which can be coupled to respondents in the backend)
- *Redis result database*: stores processing results

Hbits (MOTUS) components:

- *Receiptscanner-processor*: downloads a receipt from the sftp server, processing it (pre-OCR, OCR and post-OCR), publishes the results in the redis database and notifies the backend via rabbitmq.
- *Receiptscanner-api*: used by the backend to manage scanners (which can be coupled to respondents in the backend) and to retrieve results.

10.2.3 Flow

- Backend requests the receiptscanner-api to create a scanner entry for a respondent (if receipt scanning service is allowed and activated for that respondent)
- When the backend receives a receipt image:
  - it saves the image on the sftp server. Files are organized by scanner-uuid directories. Filenames are date and uuid based.
  - it sends an asynchronous message (*receipt.process*) on the rabbitmq bus to process the receipt image.
- Receiptscanner-processor:
  - reads the processing request message from rabbitmq
  - fetches the receipt image via sftp
  - processes the image
  - puts the results in the redis db
  - notifies the backend that processing is done via a rabbitmq asynchronous message (*receipt.processed*)
- Backend:
  - reads the notification message
  - requests the results from the receiptscanner-api
  - returns the results to the respondent when requested by the mobile/web app.

10.2.4 Receiptscanner-processor design

The receiptscanner-processor performs the main processing. It internally calls a python program to perform pre-OCR, OCR and post-OCR. The program is coded as a chain of actions in which the output of one action (block) is the input of the next block.

The internal flow is as follows:

STATBEL
Belgium in figures

D∎STATIS
wissen.nutzen.

VUB VRIJE UNIVERSITEIT BRUSSEL

h

eurostat

Figure 50: Receiptscanner-processor design internal flow
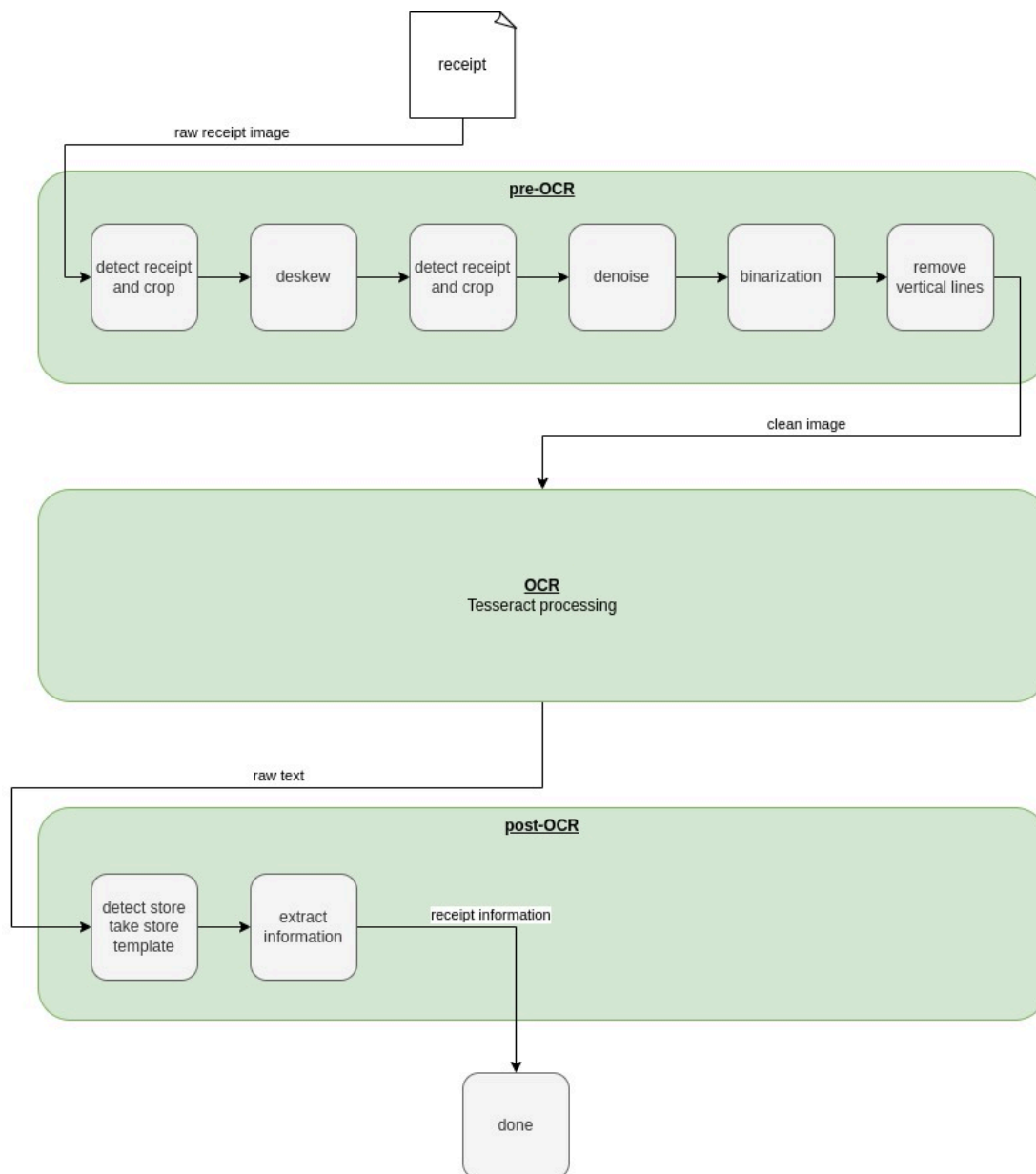
detect receipt and crop: find the receipt in the image and crop. If the receipt is slanted then there still will be quite some background in the image. Therefore:

- deskew: rotate the image to set the receipt in the image straight
- detect receipt and crop: second run to only get an image with a receipt (and no background)
- denoise: remove noise
- binarization: make black and white
- remove vertical lines: removed because sometimes they are recognized as characters. Less clutter gives better results
- detect store and take store template:
  - detect store: based on search on the raw text (output of tesseract). If store detection fails, then the flow ends unsuccessfully.
  - take store template: describes how the ticket is formatted (format of expense item rows, date, expense total etc.). Formats are defined by means of regular expressions.
- extract information: use template to extract the information.

10.3 Demo and API

10.3.1 Service deployed in a rancher cluster

In this diagram all components are deployed in a rancher cluster. Helm charts are provided in the source code deliverable.

Figure 51: Service deployed in a rancher cluster diagram
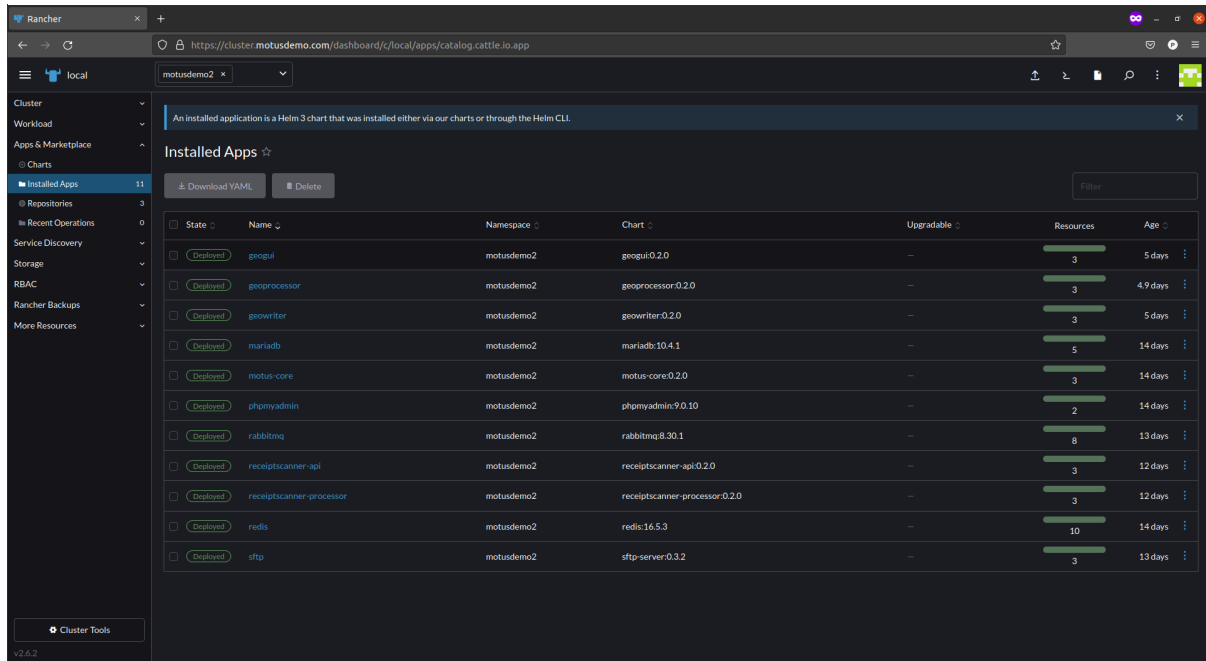
Local port forwardings are used to demonstrate the different components (redis, rabbitmq and receiptscanner-api).

10.3.2 Creating a scanner

In order to use the microservice, a scanner needs to be created via the receiptscanner-api service (insomnia screenshot is shown for demo purposes).

Figure 52: Creating a scanner
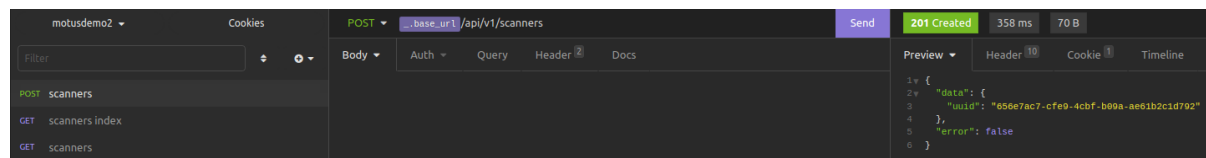
1.1.1 Scanner API

Full api is described in an openapi document *storage/api-docs/api-docs.yaml*. This document can be loaded in a browser via endpoint {base_url}/*api/documentation*.

1.2   Activating the receiptscanner-processor

In order to start the processor, an message needs to be sent to the rabbitmq broker on a *receiptscannerservice* exchange. That message will be read by a processor. The route of the message is *receipt.process*. The content of this message is a json object with the following key/value pairs:

- *scanner_uuid*: uuid of the scanner
- *receipt_uuid*: uuid of the receipt
- *timestamp*: a timestamp e.g. when the app received this image from the respondent. Format of the timestamp is *Y-m-dTH:i:sP*.
- *receipt_filename*: the image filename

The receipt processor will then fetch the file at this location of the sftp server: {scanner_uuid}/{receipt_filename}.

1.2.1   Demonstration of image processing steps

Figure 53: Image processing original image

Figure 55: Image after pre-OCR

Figure 54: Text detection by Tesseract

10.4 Obtaining the results

Results can be obtained via the receiptscanner-api as is shown here (insomnia screenshot).

Figure 56: Insomnia screenshot
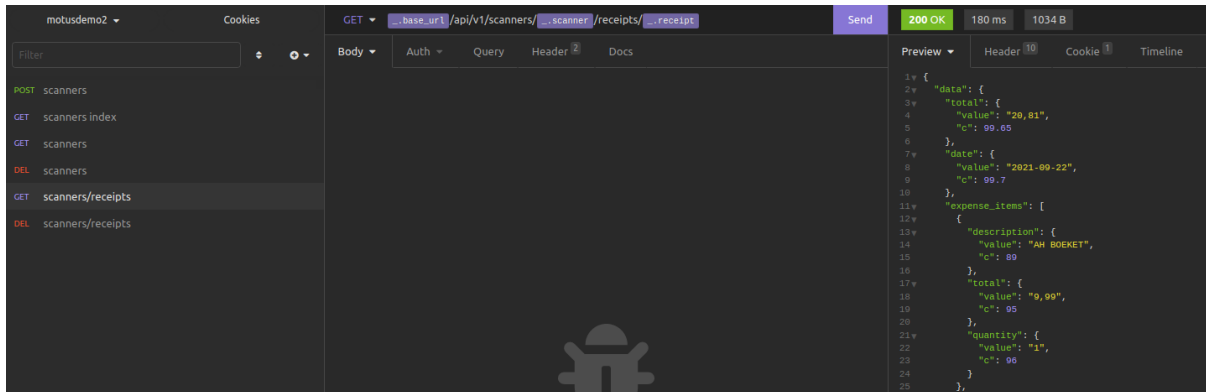
Main fields are:

- *total*: expense total,
- *date*: date found on the receipt,
- *expense items*: list of item found in the receipt.

Note that a confidence *c* is always provided as well.

1.2.2   Results in redis database

The obtained results internally come from the redis results database. The redis database is here (via the a Redis client called AnotherRedisDesktopManager):

Figure 57: Results in redis database
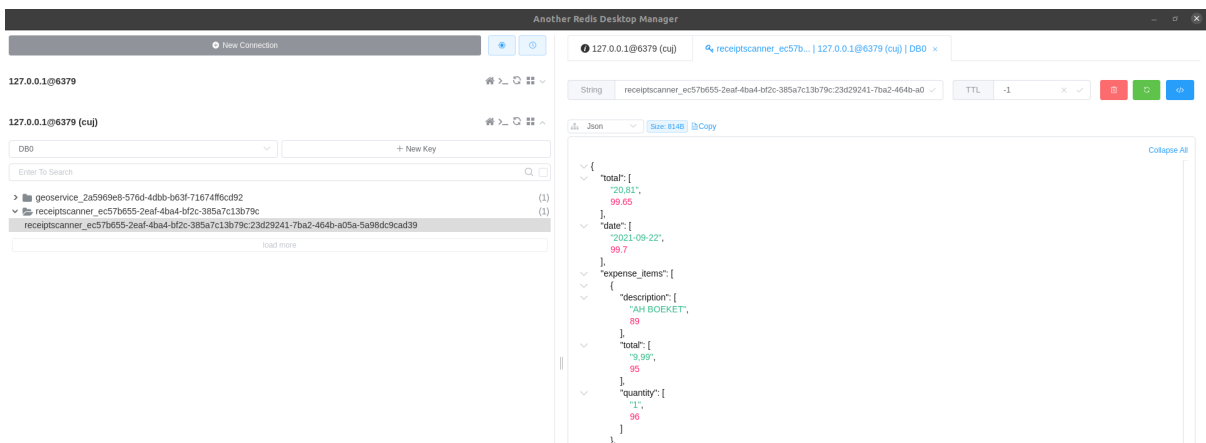
10.5 Limitations, challenges and improvements

- Pre-OCR:
  - Receipt detection algorithm: the detection of a receipt requires a high contrast between receipt and background. It is required that the respondent takes a good quality picture against dark background. Possible improvements: detection algorithm, or let user select the receipt in the mobile/web app with dots (4 dots for each corner).
  - Relatively high resolution and good image quality is required. Improvement: image processing algorithms should be adapted to work as well with low resolution images and MOTUS should provide feedback to the respondents.
- OCR:
  - Font training might improve character recognition e.g. sometimes a 0 (zero) is interchanged with an 8 (eight).
  - Exploit tesseract page segmentation modes to extract as much text as possible
- Post-OCR:
  - Currently there is a template for each store. An improvement would be a generic template so that processing is store-agnostic. This is a challenge though because each store has its own way of printing information on a receipt.
  - Store detection: if store detection fails (the store name not found in the raw text) then the process ends unsuccessfully. It might be an option to let the respondent input/choose the store, then the store template can be uniquely identified. Alternately, deep learning might be an option to detect the store.
  - Detection of receipt rows: it might be an option to let the respondent select the receipt rows on the ticket with the help of a selection box.

10.6 Installation

- unzip the receipt scanner code in a directory
- create a *.env* file (an example is provided *env.example*)
- receiptscanner-api:
  - test the api:
    - run *php artisan serve*
    - go to http://localhost:8000/api/documentation
    - alternatively, use a tool like insomnia to do api calls
  - deploy:
    - without Docker: point the web server to the *public/* directory
    - with Docker: a Dockerfile is present in the *docker/* directory. Note that there are quite some dependencies: those can be reviewed in the *docker-compose.yml* file.
- Receiptscanner-processor:
  - python dependencies: there are some libraries required to run the python code: pytesseract and python-opencv are the most important ones.
  - deploy:
    - without Docker: run *php artisan receiptprocessor:start*
    - with Docker: a Dockerfile is present in the *docker/* directory
    - helm files are available in the *helm/* directory

11 WP3 – D3.4 – R-cran package microservice

Overall goal: live demo, documentation and code via R.

11.1 Introduction

The development of R-cran package supports the data collection process, as it provides advantages to the automatization of the fieldwork via a dashboard, the cleaning of the datasets, as well as to the valorisation of data. By including this microservice the project incorporates the power of R as a statistical platform. The back-office will be able to interact with the R-microservice to support the data collection. These processes run continuously and automatically so that data is immediately available, also throughout the data collection itself.

11.2 Architecture

The R setup is different from the architecture for geoservice and receiptscanner. CPU-intensive database tasks such as R analysis are offloaded to a server which is dedicated to analysis. As such, the main database is not overloaded by complicated analysis queries.

This architecture reflects the well-known difference between 2 types of data processing systems:

- OLAP or online analytical processing, vs
- OLTP or online transaction processing.

An OLTP system is designed to do simple and fast queries to the database, while an OLAP system is designed to do slow, complicated queries to a (replication) database.

The motus server will request an R analysis and will fetch the results when available.

Figure 58: Architecture R-microservice
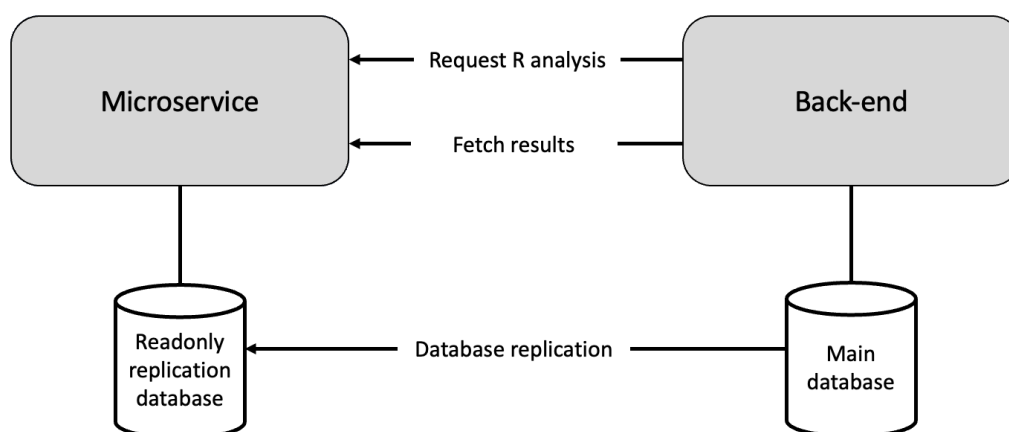
11.3 Documentation

Documentation is delivered and presented via a PowerPoint given during the WP3 CRŒSS-project, and during the Wiesbaden workshop organised for the 20th-22nd of September 2022.

The microservice reads an SQL-database that in first instance automatically employs an R-script to process Raw data in various format types. Thereafter a second R-script starts to run and has the goal to clean the data and to make checks on the quality of the collected data.
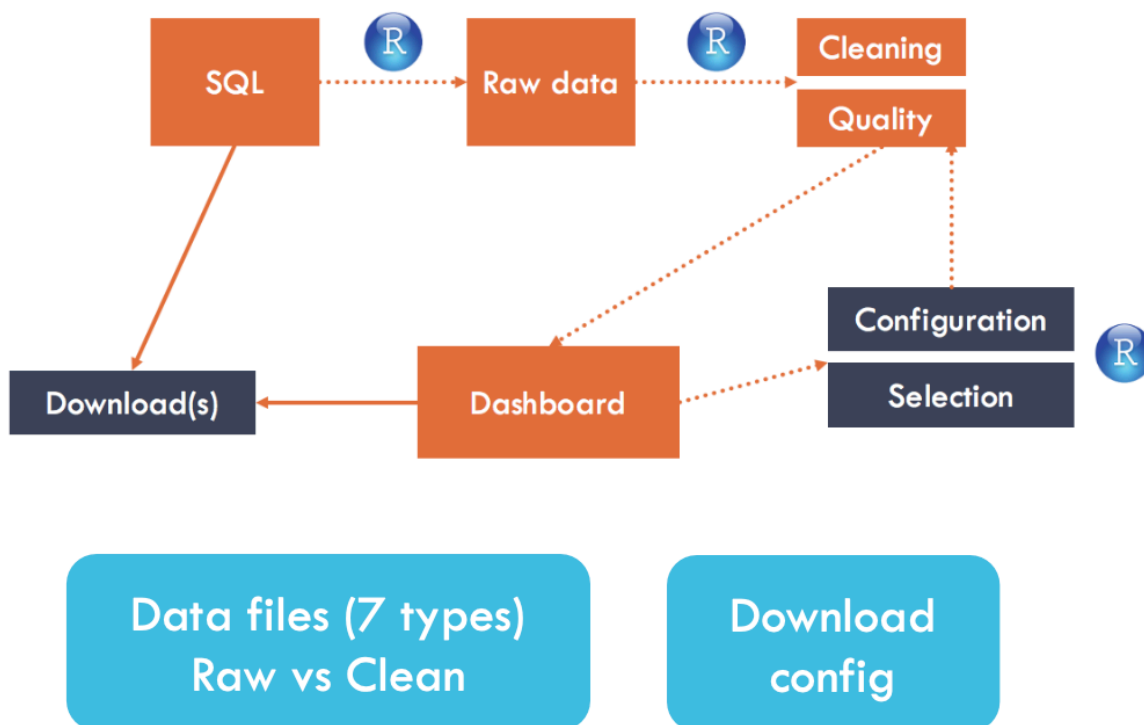
The definition of an R-script is open. Within the WP3 a choice has been made to define between:

- Diary Data Transformations, which hold
  - Database transformations like, eg. removing duplicates
  - Create extra rows, eg. adding registrations on undefined moments
  - Create extra columns, eg. grouping of levels, derived variables, episode numbering, ...
  - Week or period files
  - Day files
- Diary Quality Indicators
  - Quality checks, eg. maximum, minimum, frequency, ...
  - Checks on week level
  - Checks on day level

The output of the cleaning and quality script can be provided to a dashboard. A dashboard is typically available via a back-office functionality. In the CRŒSS project this is the MOTUS data collection platform.

The microservice also makes it possible to accept a new set of configurations based on the initial R-script (step 2). It is also possible to make a selection of the database, and reprocess this to the R-microservice. The output is than again available via the dashboard.

Via the dashboard or the microservice 7 types of data files are available, and also the configuration file can be requested as a download.

WP4 – Defining and testing of new interfaces for TUS and HBS

<u>State of play</u>

MOTUS was tested during the SOURCE™-project . Currently no application covers both TUS and HBS, nor is able to let respondents interact with smart data.

<u>Added value</u>

By means of a 1-to-n strategy and the underlying client API MOTUS was preconfigured to service both TUS and HBS surveys.

<u>Actions</u>

The main task was to evaluate new designed interfaces that connect to TUS and HBS studies and that are able to interact with data captured through microservices. Special attention had to be given to privacy-preserving screens and consent asking screens. The new design also needed to take into account the household as surveying entity.

<u>Objectives</u>

The TUS front-office (interface) has been tested in the SOURCE™-project. Remarks on the usability of the TUS application were documented. The good aspects need to be preserved, the negative remarks to be solved.

WP4 covered the existing issues while it also needed to include new innovations and developments that have an impact on the UI/UX of the front-office. The goal was to support the respondents in their data entry. A new interface was designed prior to this project. In this project the interfaces were tested and evaluated. Modifications can be made.

The specific objectives of WP4 were:

- Update/new UI and UX of the MOTUS app for TUS and HBS
- Include household level
- Include privacy-preserving screens
- Basic usability tests of the various elements

12 WP4 – D4.1 – UI/UX for TUS and HBS

> Overall goal: live demo and slides presenting the new UI/UX with introduction of the elements.

The new UI/UX was developed at the start of the project. First, a concept overview was created by hbits, based on the feedback received during the SOURCE™ project. The overview at the start of the project was accompanied by a folder with Printscreens, which can be found in annex 1, as well as the description of the screens and feedback.

The UI/UX for both TUS and HBS was tested through an iterative process. The first TUS-test was done by Statbel in December 2021 as a preparation of the national TUS data collection, which started in May 2022. Several remarks were given, MOTUS was adapted and tested by a smaller number of Statbel-colleagues until it was ready to deploy to the general population. This test was not part of this project.

The following tests were done by Destatis and Statbel and had a focus on HBS. The report on the usability tests can be found in D4.2. The final UI/UX was shown during the workshop (see D1.3).

Workshop info, agenda and PowerPoint presentations are available on the Eurostat wiki page: https://webgate.ec.europa.eu/fpfis/wikis/display/ISTLCS/Workshop2022.

The presentation of the CRŒSS project is also available on our Webpage ITNRW: "NRW connect extern" as B5460_BE_1.3_Workshop_Presentation_ENG.pdf

- The workshop presentation contains the slides presented by Statbel, hbits and Destatis, as well as a live demo about the new UI/UX.

The final presentation and live demonstration given during the workshop (see D1.3) gives the latest overview of the developments done for the CRŒSS project.

The Github directory for code on WP3 – D3.2 – Geolocation microservice and WP3 – D3.3 – Receipt scanning microservice prototype was made available for Eurostat through ZIP-files who can distribute the code as they see fit. The same procedure was applied for the WP3 – D3.4 – R-cran package microservice.

13 WP4 – D4.2 – Report on the usability tests

> Overall goal: test report, live demo, documentation and code via R.

13.1 About the usability testing

The TUS front-office (interface) of MOTUS has been tested in the SOURCE™-project (Minnen, 2020). Remarks on the usability of the TUS application were documented. The good aspects need to be preserved, the negative remarks to be solved. The goal is to support the respondents in their data collection.

A new interface was designed prior to this project. In this project the interfaces are tested and evaluated. After every test, remarks were given to the hbits team and modifications were made so the testing was an iterative process.

As a preparation to the national Belgian TUS data collection, Statbel tested the renewed MOTUS TUS UI/UX interface, that is described in D4.1.

To test the Motus design applied to HBS, both Destatis and Statbel conducted user-tests to evaluate the Motus HBS-App in regards to its usability and functionality within the German and Belgian context. These tests took place in April and September 2022. Please consider that these were not extensive pre-tests, but instead conducted internally within Destatis and Statbel, with a small sample size of eight resp. six colleagues who used the app in their daily lives for one week.

In the following, a brief description of the methodology and a summary of our main test findings are presented, including both general recommendations as well as specific propositions on how to improve the usability and functionality of the Motus HBS-app. Extensive reports, showing detailed results and suggestions for improvements, can be found in Annexes 1 (Germany) and 2 (Belgium). These reports can be considered as stand-alone descriptions of the tests in both countries.

13.2 Methodology

13.2.1 About the MOTUS-App

- MOTUS is developed by hbits and allows to track a household's expenditures for the Household Budget Survey (HBS).
- MOTUS offers both a web app, accessible via web browser, and a native mobile app, which can be installed on mobile devices. This test's focus laid on the mobile app.
- The app used for our test was still work-in-progress, meaning some content and functionalities were not yet implemented.
  - While the final app will include all tasks of the HBS, the German test version only included the central expenditure diary. The Belgian test version included a cost diary and a limited questionnaire.
  - Furthermore, among the more prominent features, the following were not or only partly implemented in both the German and Belgian test:
    - The interaction bar on the bottom of the main screen that is supposed to contain an assistant/tutorial, a search function, as well as a the "MyGroup"-functionality (allowing for household members to share expenditures, or the filter and suggestions when adding expenditures) was configured as inactive.
    - Also not part of this test version were:
      - receipt scanning,

- push notifications,
- geolocation,
- filter and suggestions function (as an aid in the process of adding expenditures),
- as well as a deviation in the general user navigation from what is planned for the final app.
  - o The beta version for the German test (April) only worked on android devices. The version used in the Belgian test (September) was also available on ios devices.
- In preparation for the German test, all functionalities, user-interface-terms and search terms were translated into German.

13.2.2 Testpersons (TP) and devices

Test 1: the German test

- Eight Destatis colleagues took part as test persons (TP).
- They varied (among others) in sex, age, smartphone use, and number of household members.
- Most colleagues were (to a certain degree) familiar with the concept of the HBS.
- We used different android-based smartphone brands (Samsung, Xiaomi, etc.) and devices. However, Apple devices did not work due to the way the beta version of the app was provided to our TP (not via app stores).
- Tablets were excluded from the test.

Test 2: the Belgian test

- Six Statbel colleagues took part as test persons (TP).
- There was some variations in sex, age, smartphone use.
- Only one participated entirely through a mobile device. Three others used both the mobile app and the web app, the two remaining test persons completed the test via the web app.
- As was the case in the German test most colleagues were (to a certain degree) familiar with the concept of the HBS.

13.2.3 Setting of the German and Belgian test

While aiming at the same target, the German and Belgian tests had different setups:

Test 1: the German test

- To gain insights for both the functionality as well as the usability, a two-step approach was followed.
- In the first step, we conducted functionality tests of certain key features, namely the search functionality as well as the "MyGroup"-tool (as far as this was possible). This allowed to test for technical issues and bugs in more detail.
- The second step consisted of the main usability test. In total eight colleagues used the app to track their daily expenditures for one week. They were asked to enter all expenditures that fall into the field phase, allowing for detailed practical insights.
- The usability test was conducted as follows:
  - o (1) TP used the app(s) on their own for one week
  - o (2) Afterwards, we conducted a focus group with all colleagues, to gain more insights into the experiences obtained by our TP and get more infos on key topics of interest

- (1) TP were asked to:
    - Download and install the app on their own smartphone
    - Report all their expenditures for one week
    - Document their experience and issues on a short feedback document
- (2) Focus group
    - We conducted a semi-standardized focus group, featuring a predetermined set of key questions, while still leaving room for an open discussion of other topics.
    - During the focus group, all TP could open the app on their device, allowing us to run through reported issues interactively.
    - The focus group lasted roughly two hours and took place right after the test week.
    - No incentives were paid to the participants.

Test 2: the Belgian test:

- The Belgian test on HBS consisted out of three tasks:

    1. Extensive testing of the cost diary over a number of days.
    2. Completing a questionnaire on the home, energy and water expenses, insurance expenses, benefits in kind and purchase of durable goods (i.e. a shortened version of the household questionnaire normally administered by an interviewer).
    3. Completing a questionnaire that gauges your experience with and opinion of the MOTUS application (installing the app, using the web application, login procedure, instructions, questioning, etc.).

- The responses to the evaluation questionnaire will be used as a starting point for a focus group interview that will delve deeper into experiences with the MOTUS software platform. An elaborate description, partially based on bilateral exchanges with the six respondents is already in the annex, however.

13.3 Summary

Overall, the MOTUS app is quite advanced in its development. The general functionality and appearance are rated positively overall. However, the overall impression of the app for daily tracking of expenditures is mixed. While the app works mostly well for adding common expenditures, the usability test of Destatis unveiled several areas that clearly need to be improved upon.

Most critically, test persons (TP) considered the product search function dissatisfying, as fitting categories could not be found for a significant variety of products. Inserting long receipts was considered troublesome, as many small positions had to be typed in manually. Entering details to an expenditure turned out to be error-prone, namely the distinction between groceries per unit (e.g. apples) and quantity of the product. Generally, TP missed additional information about how certain functionalities work and what a respondent is supposed to add where. Moreover, navigating through the app was not considered as intuitive by some TP, due to certain elements of the layout and the button display.

14 WP4 – D4.3 – User manuals

Overall goal: user manuals in store, in app, on website, assistant.

A comprehensive overview about MOTUS, who is behind it and how it works can be found online: https://www.MOTUSresearch.io/

Figure 60: MOTUSresearch.io
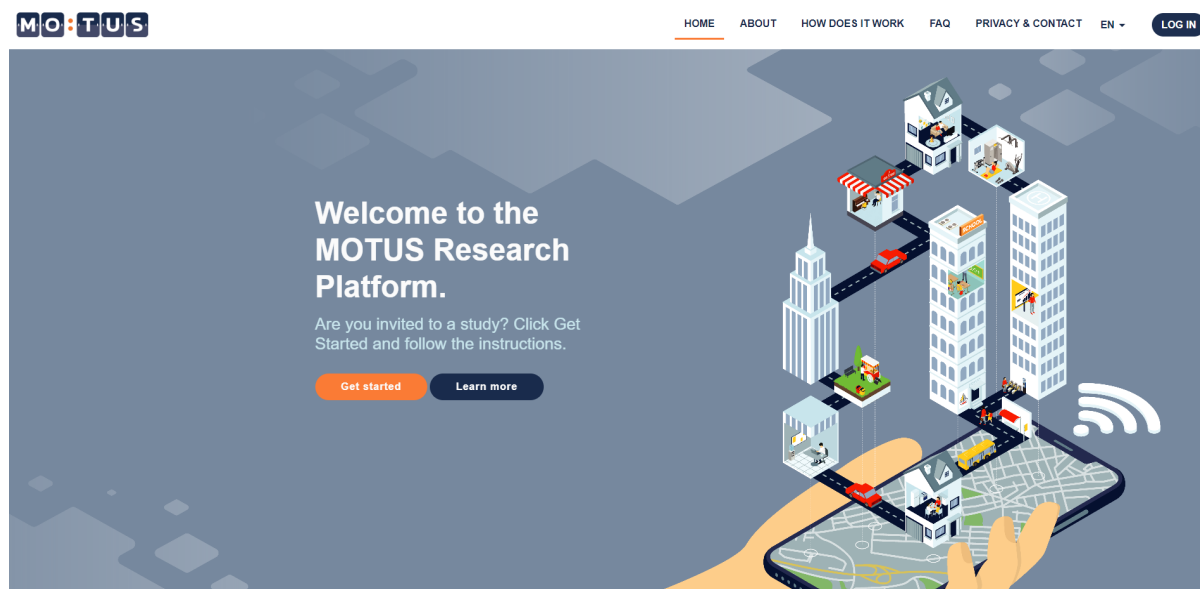
The website does not only provide user manuals, but also:

- the FAQs,
- privacy-preserving screens,
- and contact information.

The same information can be found in the:

App store: MOTUS Discovery:

https://apps.apple.com/be/app/motus-discovery/id956934466

Play store: MOTUS Discovery:

https://play.google.com/store/apps/details?id=be.byteworks.motus

And in the MOTUS-app itself. All information is available in English (default language), Dutch and French and certain parts in German.

WP5 – Roll-out research data collection platform

<u>State of play</u>

Eurostat as well as other international bodies encourage software developers/owners to document their tools according to the CSPA-catalogue. The CSPA has 3 levels: a conceptual level, a logical level and a implementation level. Extra attention is given to the criteria of reusability, shareability and comparability. MOTUS has been already documented on the conceptual and logical level.

<u>Added value</u>

Through the modular setup and the availability of both a front- and back-office, MOTUS is seen as a shareable tool. During the CRŒSS project this shareability is further extended: by including HBS (WP2), by including new functionalities (WP3) and by including a new UI and UX design (WP4).

By including a roll-out strategy in WP5 the aim is to improve the shareability of MOTUS within the ESS. Also taking into account today's security and privacy requirements.

<u>Actions</u>

The project has the objective to develop a prototype where the installation of the MOTUS-platform is implemented through containers. This way MOTUS is easy to share and can be integrated into the NSIs own production line. The MOTUS-code runs on the 'application tier' of an NSI by means of the Docker technology. In this way the development team of MOTUS can further develop/update and maintain the platform while at the same time comparability and shareability are maximized.

In essence the container technology provides four pillars that are essential for a reusable, shareable, flexible and maintainable ESS platform. These pillars are:

1. Deployment, meaning: remote/cloud deployment, fast installation, increased error reduction and an organised validation control, …
2. Scalability, meaning: multiple databases, replication, many/large projects in size and numbers of respondents at the same time without loss of performance, strong improvement of both analysis and transaction functions to accept a large number of short online transactions, …
3. High-Availability, meaning: database replication and replication of front-end connectivity to deal with system failures, …
4. Privacy & security, meaning: strict separation of databases, role management, …

<u>Objectives</u>

The second overall goal was to make the MOTUS data collection platform for TUS and HBS easier shareable amongst NSIs in order to support them in their data collection strategy. In the SOURCE™-project 4 different roll-out strategies were defined and evaluated. The conclusion was to further investigate the container technology. This way an interpretation was given to the Governance Guide that Eurostat is defining together with the joint TF TUS and HBS.

The specific objectives of WP5 are:

- Define an efficient roll-out strategy to share MOTUS.
- Define the requirements to host MOTUS and integrate it to an existing production line.

15 WP5 – D5.1 – Document database technology pillars

> Overall goal: documentation of the choices made to define the different containers.

15.1 Introduction

One of the goals of the CRŒSS project is to make the MOTUS data collection platform easier to share amongst NSIs in order to support their data collection strategy. This document describes the architectural and technological choices being made to achieve this goal. It does so by defining a container-based roll-out strategy. This container-based solution was evaluated as the best rated option between 4 possible solutions in the SOURCE™ project report, and allows the MOTUS platform to evolve to an ESS-platform.

15.2 Requirements

For MOTUS to evolve to an ESS-platform, it needs to be able to fulfill the following requirements:

- Easy deployment: fast and easy deployment of the MOTUS software stack at different locations and/or servers. This avoids repetitive, time-consuming and possibly error-prone installation tasks, and enables the platform to setup a CI/CD pipeline (continuous integration / continuous deployment).
- Scalability: the MOTUS platform needs to be able to handle an increasing number of respondents. More respondents means more concurrent requests to the system. In addition, it requires more time to perform analytic operations to retrieve data and statistics. Scalability is important for both MOTUS core and microservices. E.g. the receipt scanner must be designed in such a way that it can analyze shopping tickets in parallel.
- High availability: the system needs to be designed in such a way that it is ready for high availability, which means trying to maximize the service up-time.
- Keeping privacy: architectural/design decisions need to take privacy into considerations, especially in the context of microservices. The MOTUS core contains sensitive data and the microservices, which can be regarded as satellites around the core, should not have direct access to it.

15.3 Solutions

15.3.1 Easy deployment

Docker as a building block

In order to facilitate deployment, an ecosystem of components will be introduced which – together – realize the MOTUS platform. Each component is a container and the chosen technology is Docker. The benefit of a Docker container is separation of the software application from the underlying infrastructure. Essentially, a Docker application is able to run as long as the infrastructure is able to start Docker containers. Dependencies that are required to run the application are all part of the software stack in the container itself.

Given the Docker container technology, deployment of the MOTUS ecosystem only requires:

- an infrastructure to run Docker containers, and
- a correct MOTUS configuration.

Docker platform

Several platforms exists on which Docker containers can be deployed. Examples are docker-composer, Docker swarm, Kubernetes and various cloud-based platforms. In addition, in order to simplify container orchestration (provisioning, deployment, management, scaling, networking etc.), platforms can be extended with extra tooling. For example, OpenShift and Rancher build upon the Kubernetes platform to make some orchestration tasks easier.

The MOTUS platform does not enforce a specific Docker deployment or orchestration platform. Hbits will use Rancher for prototyping. Rancher is on the one hand a Kubernetes distribution, on the other hand a web-based UI dashboard which is able to connect to several Kubernetes clusters. The arguments for choosing Rancher are:

- based on well-known and proven Kubernetes cluster technology,
- easy to install,
- open source with a strong community.

MOTUS ecosystem of Docker containers

The MOTUS platform consists of the following components:

- Presentation tier
  - o Front office: web app or mobile app. These are the user interfaces for the respondents.
  - o Back office: web app for researchers and admins.
- Application tier
  - o Back-end application: application logic
- Data tier
  - o Back-end database(s)
- Microservices: satellites around the back-end. Loosely coupled with the back-end. Might implement all tiers.

Mapping to Docker containers is proposed as follows:

- mobile app: not applicable
- the back office, back-end and database components each become individual containers
- microservices have a more complicated structure as was discussed in the Geolocation microservice architecture document (already delivered in the project)

Figure 61: Overview of basic setup
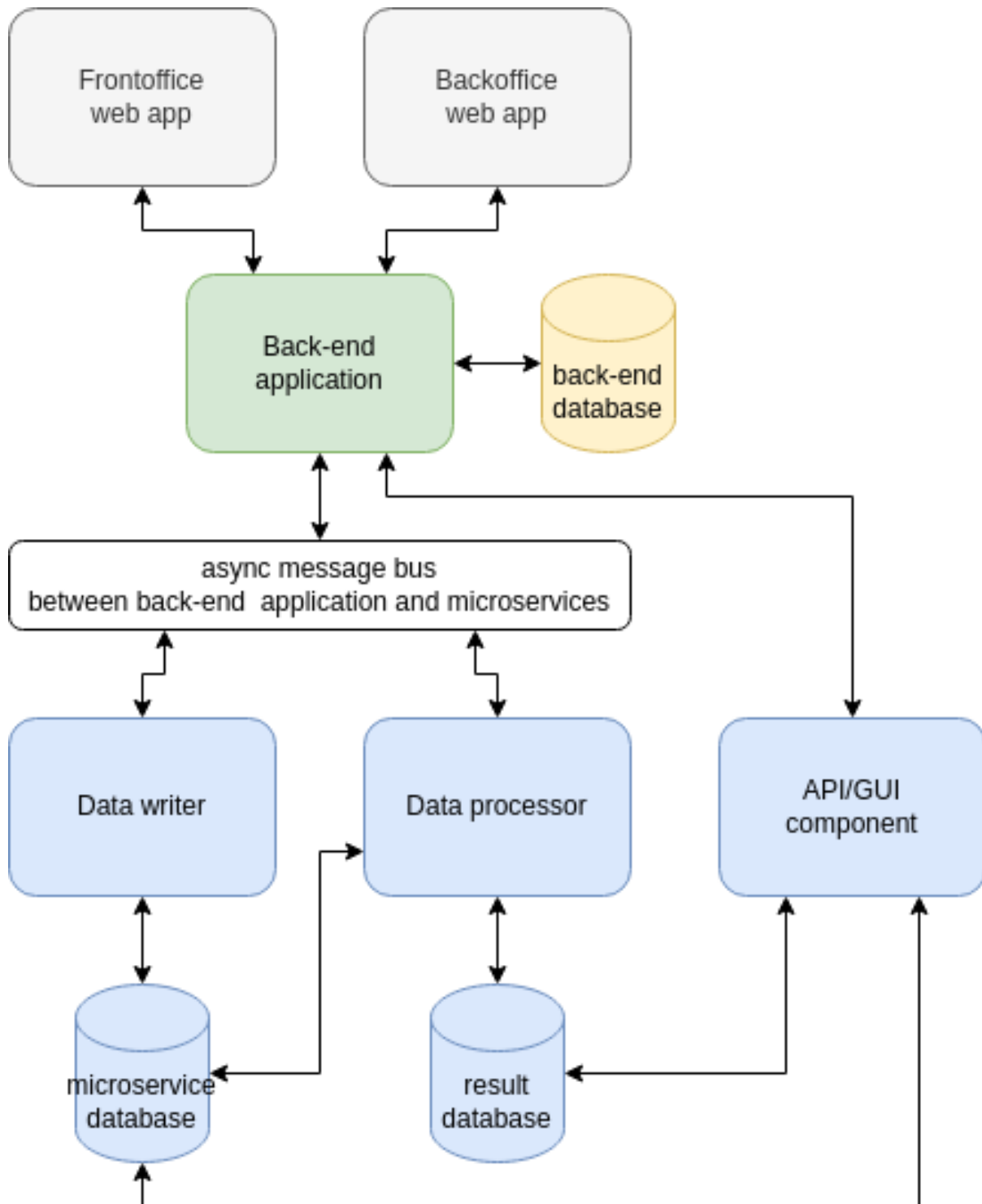
The blue boxes represent a single microservice. All boxes and database representations are different Docker containers.

The setup can be more complicated when high-availability and scalability requirements are taken into account.

Note that it is not a strict requirement that all components are deployed as a Docker container. For example, one might keep a non-Dockerized Oracle database instead of a Dockerized MariaDB. It is up to the NSI to decide how to do the deployment and which requirements are applicable.

Development pipeline

In order to support Continuous Integration and Continuous Delivery (CI/CD), a Jenkins server will be used in combination with a Docker registry. The Jenkins server enables automated building, testing and deploying of artifacts (e.g. Docker containers). It does so by means of a configurable pipeline of actions.

15.3.2   Scalability

The scalability requirements can be solved in different ways:

- Because of easy deployment, the MOTUS platform can be (re)deployed per project or - within one project - per geographical region. As such, no single server need to do all.
- Replication of the back-end database, OLTP (fast transaction processing) vs OLAP (slow analytical processing): a database for end-user input (very fast, small transactions) which is independent from a database for researchers (slow and complex database queries). More concretely, the main database could be (read-only) replicated. Database analytics could then be performed on the replicated slave database thereby offloading the master database.
- Scalability within a microservice: this is discussed in depth in the delivered Geolocation microservice architecture document.

15.3.3   High availability

The 3-tier MOTUS platform architecture enables high availability use cases.

For example, a single instance of a MariaDB database could be replaced by a Galera cluster. This technology allows for a multi-master database setup.

As high availability solutions are typically complex and more difficult to setup and maintain, for each different project a study is required to see whether such a solution is worth the extra cost and effort.

15.3.4   Privacy

The sensitive data of respondents is stored in the back-end database (and its replications). Multiple actions are required and several design choices have been made to protect this data:

- The infrastructure which 'surrounds' the database needs to adapt sound security policies in order to protect the servers from unwanted access. E.g. the platform could be deployed in a ISO/IEC 27001 server environment.
- As mentioned in the project, the MOTUS application will be subject to PEN tests.
- Separation of user data depending on geographical region (risk mitigation).
- Microservices do not have direct access to the back-end database:
  - data access from microservice to main database is under full control of the back-end application
  - the concept of a respondent is not known to the microservices. Instead, as explained in the Geolocation microservice architecture deliverable, microservices work with a more abstract concepts like trackers (geotracker) or scanners (receipt scanning).

16 WP5 – D5.2 – Installation report of the MOTUS containers

> Overall goal: report on the past experiences of the native installation, installation of the MOTUS containers and listing of the necessary requirements.

16.1 Introduction

In order to facilitate deployment (see D5.1 Document database technology pillars), MOTUS delivers a collection of Docker containers. The main benefit of Docker containers is the decoupling of the application and its dependencies from the underlying infrastructure. Additional benefits are: more lightweight than virtual machines, easier to scale and the availability of Docker platforms. A Docker platform is able to run, manage and coordinate Docker containers. Hbits choose Rancher because it is open-source, Linux-OS independent and its UI supports multiple clusters. Other alternatively can be equally good, Hbits delivers Docker containers which are not strongly coupled to a specific Docker platform.

This report shows how an installation can be done with Rancher and its software management tool Helm. The configuration files can be handed out on request

16.2 Report on the past experiences of the native installation

In preparation for the German TUS conducted in 2022, a Proof of Concept (PoC) of a native MOTUS installation was piloted, in which a Regional Statistical Institute (RSI) hosted all its components. Running this PoC the results of the SOURCE™ Eurostat project have been a very helpful basis. The PoC resulted positive in a productive use of a native installation for the German TUS wave 2022. This native installation was specifically customized towards German needs, including:

- A customized version of MOTUS back-office (MOTUS builder)
- Exclusive German mobile apps for Android and iOS published via the App Store Accounts of the RSI
- An exclusive German web frontend published under a custom domain by the RSI
- Configuration of all components and interfaces to allow for a stable and reliable operating environment in the high protection IT infrastructure hosted by the RSI

Preparing the native installation was a lot of effort for both parties involved: hbits as provider of the software and support partner for the hosting service provider and the RSI as the hosting service provider. It turns out, that IT infrastructures are differ and even small differences from one hosting partner to another can lead to issues when setting up an operating environment, even if – at first glance – the requirements seem to be similar. The task of setting up a native installation are further impaired by given data protection laws and IT security requirements that hinder or prohibit the free flow of useful information (e.g. log files) between partners. These problems enlarge when partners work in different places.

Nevertheless, the native installation used in production of the German TUS must be seen as a first positive outcome of the projects initiated by Eurostat. The MOTUS platform currently proves its possibilities conducting an official statistical survey in a member state.

Nonetheless, the SOURCE™ has already rated MOTUS containerisation as the best option to evolve MOTUS to an ESS-platform. The experiences made and described above again underline this finding and give yet another reason to go forward with the installation via MOTUS containerisation described in the following sections.

16.3 Installation report of the MOTUS containers

16.3.1 Requirements

The requirements for deploying MOTUS depend on the project. The minimum requirement is a platform to manage and run Docker containers. The complexity of Docker platforms differs widely: from bare-bones to complex clusters. Some examples are:

- Docker daemon: basic, mainly manual control such as starting/stopping container
- docker-compose: allows for describing the setup in a yaml file, setup/breakdown is done with a single command
- cluster solutions: powerful but complex. A well-known cluster solution is Kubernetes, which can be extended with docker orchestration software (auto-configuration, management, scaling etc.) such as Rancher and OpenShift

For MOTUS, a cluster solution is advised as this make use of the full power of a Docker deployment:

- auto-scaling of microservice processors
- auto-restart of Docker containers at exit
- control of resources
- load-balancing possibilities

16.3.2 Overview installation

Diagram

Figure 62: Overview installation diagram
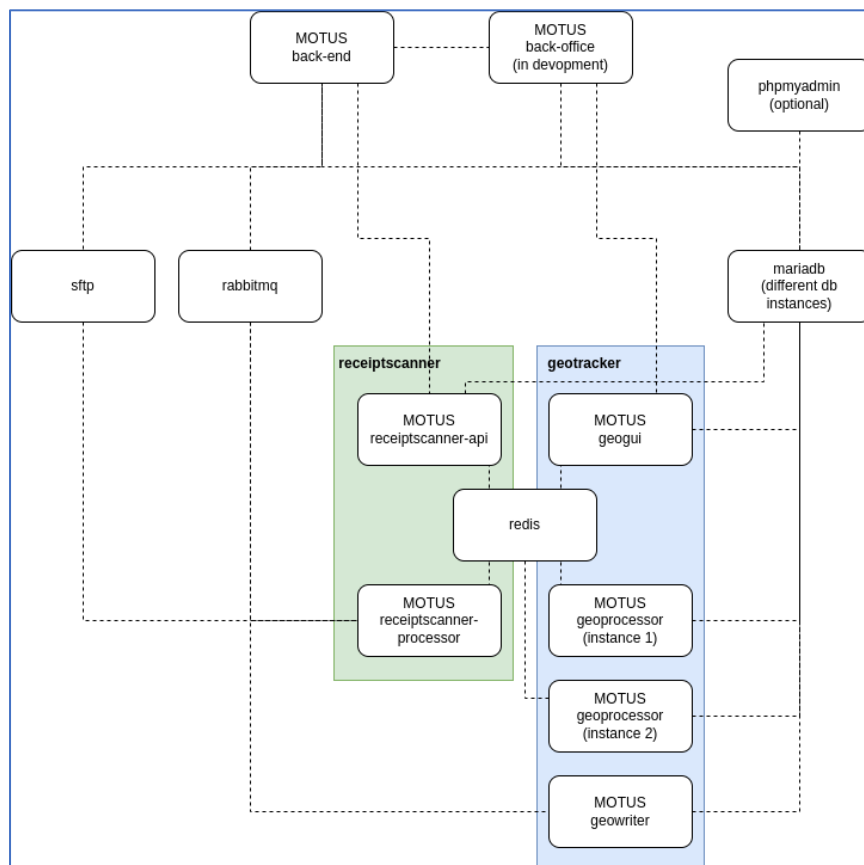
MOTUS containers

- MOTUS Back-end
  - MOTUS-core: application logic (dependencies: mariadb, sftp)
- MOTUS Microservices
  - geotracker
    - geogui: gui and api, used by MOTUS-core to manage trackers and retrieve results (dependencies: mariadb, redis)
    - geoprocessor: processes the geo tracking points (dependencies: mariadb, redis, rabbitmq). In this installation, for the sake of demonstrating horizontal scalability, 2 instances have been created.
    - geowriter: writes tracking points to the geotracker db (dependencies: mariadb, rabbitmq)
  - receiptscanner
    - receiptscanner-api: api used by MOTUS-core to manage scanners and retrieve results (dependencies: mariadb, redis)
    - receiptscanner-processor: takes a shopping ticket and processes it (dependencies: mariadb, redis, rabbitmq)
- MOTUS Backoffice (in development)

16.3.3 Non-MOTUS containers

- sftp: transfer of image files from MOTUS back-end to receiptscanner microservice
- rabbitmq: asynchronous message broker between MOTUS backend and microservices
- redis: key/value database, 1 redis instance is shared between microservices. It is possible to instantiate multiple redis instances if needed so.
- mariadb/phpmyadmin: 1 mariadb instance that has several separate databases: one for the backend, one for receiptscanner, one for the geotracker and also a limited product database.

16.4 Installation of MOTUS components on Rancher with Helm

16.4.1 MOTUS-core

Prerequisites before installation

Storage

- A persistent volume e.g. <namespace>-MOTUS-core-storage-pv. In a production environment a network storage solution such as GlusterFS or CephFS should be used but in this test setup a HostPath to a local directory was chosen. Host permission to the storage directory must be www-data or 33:33 (userid:groupid). Furthermore, the laravel storage directories need to be present (e.g. email_layouts, app, app/public, logs storage/framework/{cache,testing,sessions,views}).
- A persistent volume claim for namespace <namespace> with the exact name: *MOTUS-core-storage-pvc*. The name is important, as it is referenced (hardcoded) in the MOTUS-core helm chart. This persistence volume claim needs to reference the earlier created persistent volume <namespace>-MOTUS-core-storage-pv.

Configmap

The MOTUS-core requires configuration for the laravel environment (.env file). This confiiguration is specific for each deployment. In this prototype setup the config map was created with the Rancher UI. In a production environment, a scripted method might be preferred with a yaml file (see for example this link: https://stackoverflow.com/questions/57219325/how-to-create-a-config-map-from-a-file-in-helm).

The following is needed:

- A config map for namespace <namespace> with the name *MOTUS-core-env* (exact name as it is referenced by the MOTUS-core Helm chart).
- Content: read the 'env' file from a file. Data is 'env', value is the content from the file.
- To update the configmap: https://stackoverflow.com/questions/49989943/kubectl-update-configmap , see Steve Coopers answer.

Installation

Using helm:

helm install --namespace <namespace> -f values.yaml MOTUS-core ./MOTUS-core/

The values.yaml only contains the following:

image:
 repository: <IP:PORT of docker registry>/MOTUS-core
 tag: "<VERSION>"

Expose the service

In order to make the service available with a domain name and via https, the following steps are needed:

- In case of letsencrypt certificate: an issuer and certificate is needed
- cert-manager is required for managing the certificate
- An ingress for MOTUS-core in namespace <namespace>:
    - request host is the domainname e.g. opq.xyz.be
    - prefix is /
    - target service is MOTUS-core port 80
    - add certificate: host is again opq.xyz.be
    - creating/updating the ingress could be done in the Rancher UI, as is shown here, or via a yaml file (preferred method for production systems)

16.4.2   Receiptscanner microservice

Introduction

2 types of containers:

- receiptscanner-api: only 1 instance, dependencies are redis and mariadb
- receiptscanner-processor: several instances are possible, dependencies are redis, mariadb and rabbitmq

Prerequisites before installation

Storage for receiptscanner-api

- Similar to MOTUS-core except for naming: use *receiptscanner-api-storage-pvc*

ConfigMap for receiptscanner-api

- Similar to MOTUS-core except for naming: use *receiptscanner-api-env*

Storage for receiptscanner-processor

- Similar to MOTUS-core except for naming: use *receiptscanner-processor-storage-pvc*

ConfigMap for receiptscanner-processor

- Similar to MOTUS-core except for naming: use receiptscanner-processor-env

Installation

Receiptscanner-api

Using helm:

helm install --namespace <namespace> f values.yaml receiptscanner-api ./receiptscanner-api/

The values.yaml only contains the following:

image:
 repository: <IP:PORT of docker registry>/receiptscanner-api
 tag: "<VERSION>"

Receiptscanner-processor

Same as receiptscanner-api except use receiptscanner-processor.

Note: there is a dependency with rabbitmq. The processor expects a rabbitmq vhost *receiptscanner*.

16.4.3 Geotracker microservice

Introduction

3 types of containers:

geogui: only 1 instance, dependencies are redis and mariadb

geoprocessor: several instances are possible, dependencies are redis, mariadb and rabbitmq

geowriter: only 1 instance, dependencies are mariadb and rabbitmq

Prerequisites before installation

Storage for geogui

- Similar to MOTUS-core except for naming: use *geogui-storage-pvc*

ConfigMap for geogui

- Similar to MOTUS-core except for naming: use *geogui-env*

Storage for geoprocessor

- Similar to MOTUS-core except for naming: use *geoprocessor-storage-pvc*

ConfigMap for geoprocessor

- Similar to MOTUS-core except for naming: use geoprocessor-env

Storage for geowriter

- Similar to MOTUS-core except for naming: use *geowriter-storage-pvc*

ConfigMap for geowriter

- Similar to MOTUS-core except for naming: use geowriter-env

Installation

GeoGui

Using helm:

helm install --namespace <namespace> f values.yaml geogui geogui/

The values.yaml only contains the following:

```
image:
 repository: <IP:PORT of docker registry>/geogui
 tag: "<VERSION>"
```

GeoProcessor

Same as GeoGui except use GeoProcessor.

Note: there is a dependency with rabbitmq. The processor expects a rabbitmq vhost *geoservice*.

GeoWriter

Same as GeoProcessor except use GeoWriter.

Note: there is a dependency with rabbitmq. The processor expects a rabbitmq vhost *geoservice*.

16.5 Installation of non-MOTUS components on Rancher with Helm

16.5.1 *Mariadb*

Prerequisites

- Persistent volume
- persistent volume claim in correct namespace
- a values.yaml files with values that need to be merged in the helm chart
    - authentication: auth: rootPassword, database, username, password
    - storage: primary: persistence: existingClaim

Helm

helm repo add bitnami https://charts.bitnami.com/bitnami

helm install mariadb -f mariadb-values.yaml --namespace <NAMESPACE> bitnami/mariadb

16.5.2   PhpMyAdmin

Helm

helm repo add bitnami https://charts.bitnami.com/bitnami

helm install phpmyadmin -f values.yaml --namespace <namespace> bitnami/phpmyadmin

Accessing the db with PhpMyAdmin from a local machine

Given the *kube_config_cluster.yml* of the cluster, run:

- kubectl port-forward --namespace MOTUSdemo2 svc/phpmyadmin <PORT>:80
- now you can use phpmyadmin from a browser at http://localhost:<PORT>

16.5.3   Redis

Helm

helm repo add bitnami https://charts.bitnami.com/bitnami
helm install redis -f values.yaml --namespace <namespace> bitnami/redis

16.6   Rabbitmq

16.6.1   Helm

helm repo add bitnami https://charts.bitnami.com/bitnami

helm install rabbitmq -f values.yaml --namespace <namespace> bitnami/rabbitmq

16.6.2   Sftp

Prerequisites

- A persistent volume + claim
- Clone repo https://github.com/openvnf/sftp-server

Helm

helm install sftp -f values.yaml --namespace <namespace> sftp-server/

16.7 Screenshots

16.7.1 Deployments view

The installation was done in a separate namespace *MOTUSdemo2*. All components are deployed and in a working state (green bar at the right side).

Figure 63: Deployment view
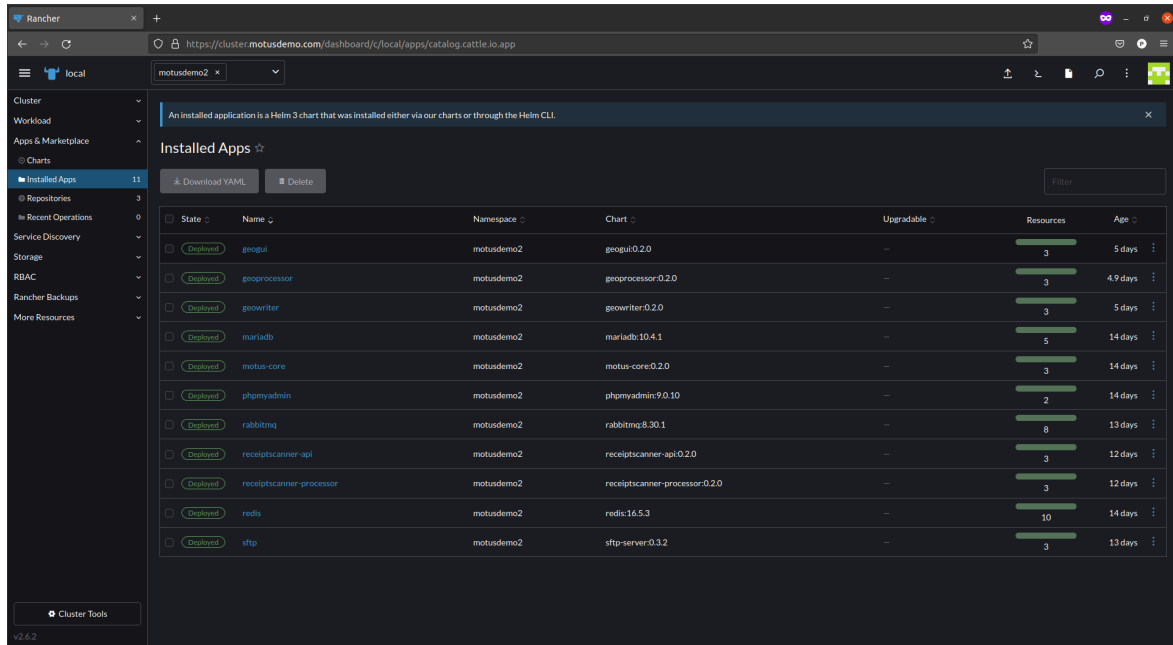
16.7.2 Pods view

Runtime overview of the installation. Note that - in order to demonstrate horizontal scaling - the geoprocessor has 2 running instances.

Figure 64: Pods view
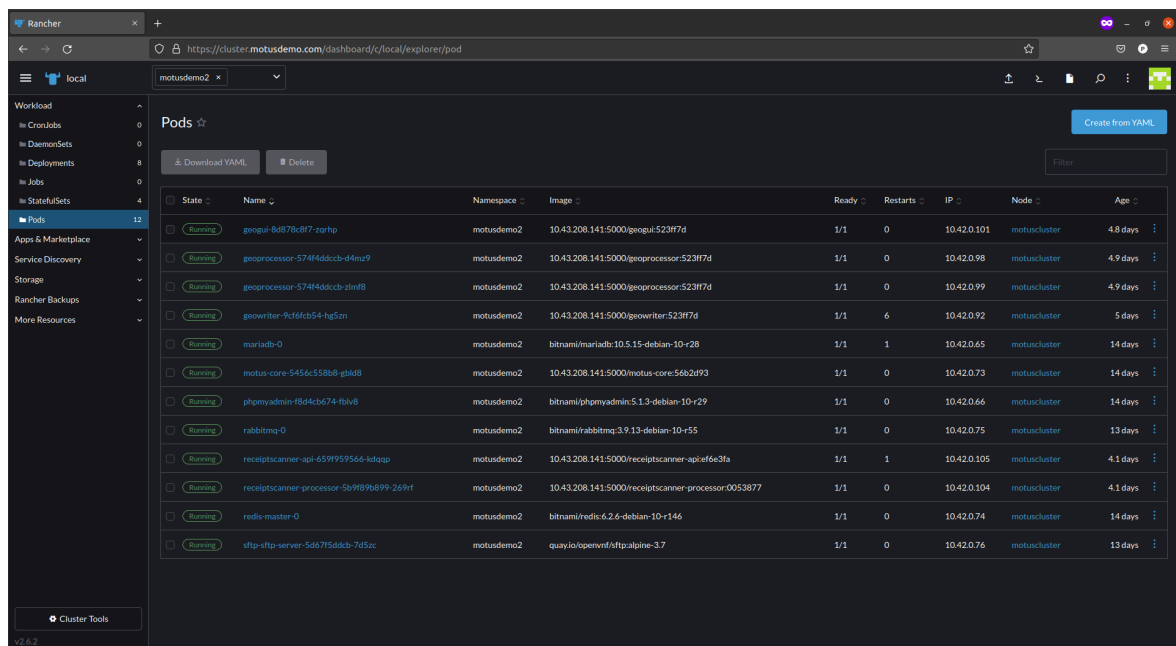
16.7.3   Ingress view

There is only one ingress, which is to MOTUS backend. The ingress is deployed with a letsencrypt certificate with the help of cert-manager.

Figure 65: Ingress view
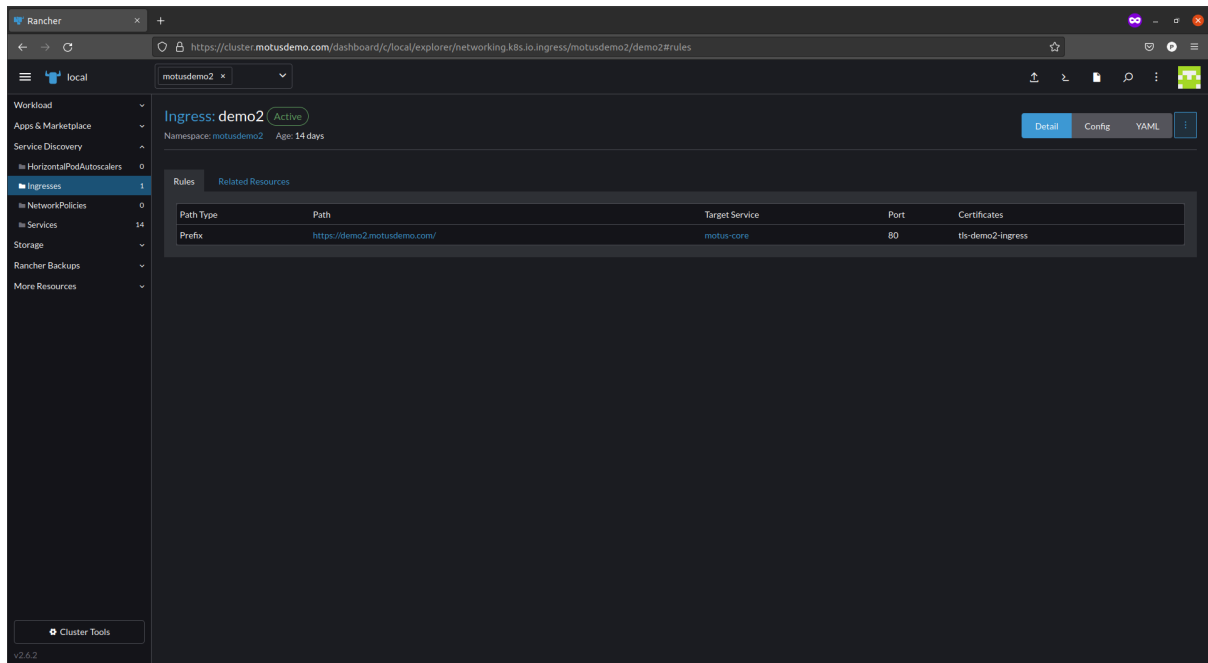
16.7.4   Overview config maps
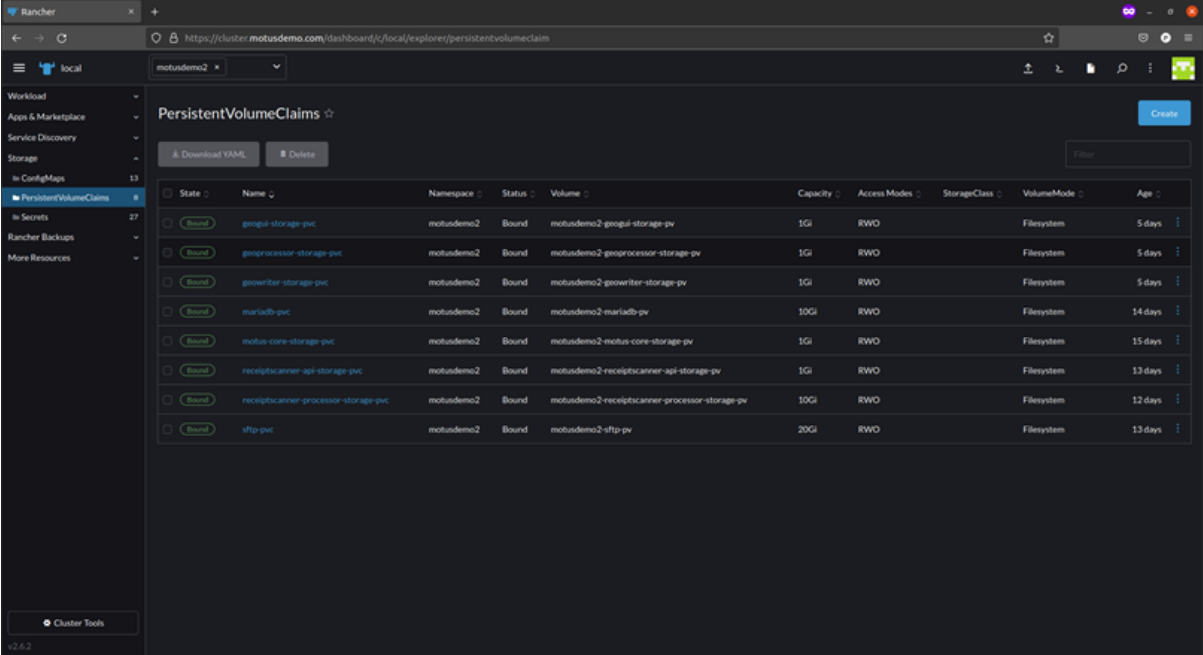
All configuration needed to run the containers.

Figure 66: Overview config maps
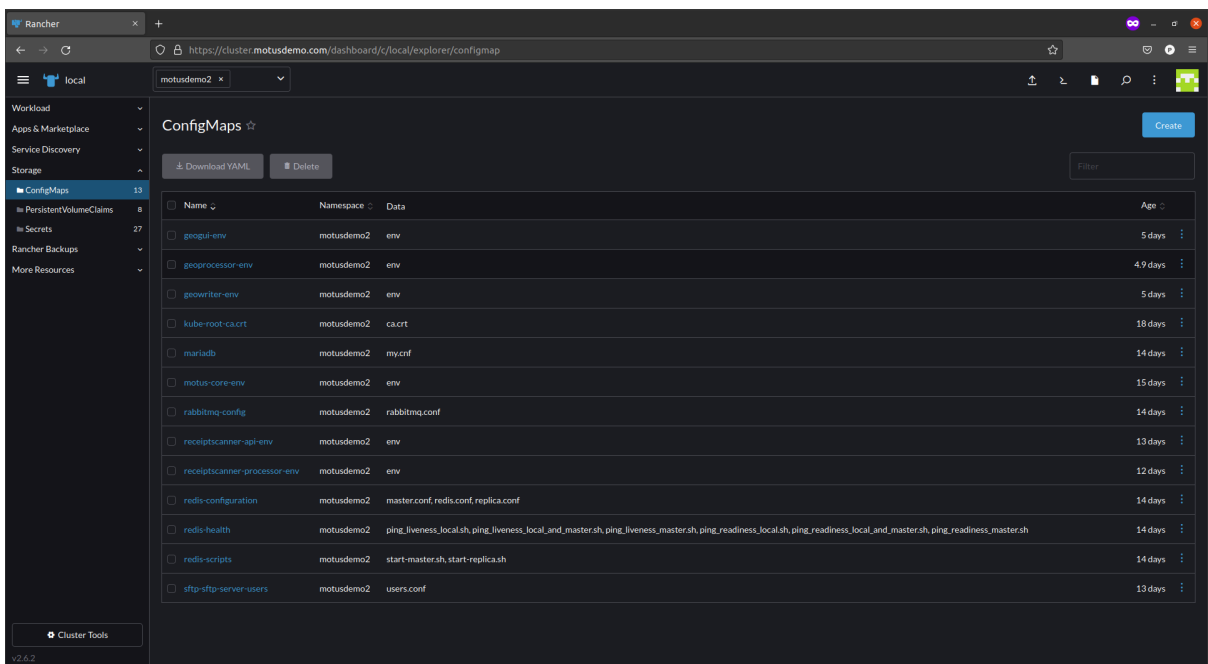
16.7.5 Overview persistent volume claims

Figure 67: Overview persistent volume claims

17 WP5 – D5.3 – Report on vulnerability tests

Overall goal: report on vulnerability tests.

In essence the container technology provides four pillars that are essential for a reusable, shareable, flexible and maintainable ESS platform. One of these pillars is privacy & security, meaning: strict separation of databases, role management, …

In relation to this, more and more organizations and software solutions have to prove their credibility via certificates. This is increasingly being asked from regulators, clients and the public. Such a particular certificate is the ISO27001. It is "the international standard that lays out the specifications for implementing an ISMS (information security management system.)". However, the ISO in itself is fairly broad. Therefore, we will within this project focus on the parts that are related to the development of tools and the quality assurance that need to be incorporated. In other words, this project will perform a vulnerability check of the MOTUS data collection platform to determine the level of protection for the collected data.

17.1 Concept of vulnerability

The concept of vulnerability within software systems is understood as a risk to run into security problems, for the system itself but also for the participant. Mapping the vulnerability then can go as broad as to run penetration tests, to perform a Source Code Analysis, to write a Data Protection Plan or to have a Privacy Declaration in place.

As being described earlier, MOTUS can be used as a service where the installation is on the premises of the Vrije Universiteit Brussel and hbits CV as its Spin-off or can be installed (natively or via virtual containers) on the premises of a responsible institution (NSI, university, company…).

This Deliverable first presents the Privacy Declaration of MOTUS and the Data Protection Plan of hbits CV. Subsequently it will document upon the penetration and source code analysis that was conducted in order to evaluate the security risks in the software system, and to detect security loopholes. Solving these risks reduces the probability of a threat.

Besides this hbits CV give its developers the opportunity to train themselves in secure coding and performs reviews of security architecture and source code. A continues review is necessary as a software platform always grows in size and in complexity, certainly when making use of third-party applications (see microservices).

17.2 Privacy Declaration of MOTUS

The Privacy Declaration of MOTUS is available online via https://www.motusresearch.io/privacy/full/ and via de different application stores.

With MOTUS we ask users (the "Respondents") to take part to a study (the "Study"). Studies include the completion of questionnaires, diaries and the inclusion of sensor data. The latter includes location information. A study has an assigned duration in advance and is voluntary. Studies can offer personal feedback to the respondent.

Studies are designed and executed by the Vrije Universiteit Brussel (VUB-Belgium) and hbits CV as spin-off of the Research Group TOR of the VUB.

Your privacy is very important to us. Before we get to the details, check out our Privacy Label to see a summary of our privacy practices. Below the Privacy Label the full policy can be found.

17.2.1 Privacy lable

Table 12: Privacy lable

Collection of data	
- Do we manage, store and process your Personal data during the study?	Yes
- Do we manage, store and process your Personal data after the study?	No
- Do we use external plugins to collect, manage, store and process Study data?	No
- Do we collect, manage, store and process User data?	Yes
Tracking	
- Do we track your device location when you use MOTUS?	Yes, with your consent
- Do we track your device location when you are not using MOTUS actively?	Yes, with your consent
- Do we use cookies?	Yes, but only first party cookies
Processing of data	
- Do we pass on your Personal data to research partners?	No
- Do we pass on your Personal data to commercial companies?	No
- Do we pass on your Study data to research partners?	Yes
- Do we pass on your Study data to commercial companies?	No
- Do we pass on your Sensor data to research partners?	No
- Do we pass on your Sensor data to commercial companies?	No
- Do we pass on your User data to research partners?	No
- Do we pass on your User data to commercial companies?	No
- Do we process Study data to aggregated information?	Yes
- Do we share or sell aggregate information?	Yes
- Do we use your Study data to apply for research funding?	Yes
Privacy and Control	
- Can someone else see your data in MOTUS?	No
- Can you control the privacy settings in MOTUS?	Yes
- Can you exercise your rights under GDPR?	Yes
- Can you make a complaint?	Yes

17.2.2 Privacy policy

The MOTUS privacy policy (the "Privacy Policy") concerns the MOTUS platform developed by the Research Group TOR of the Vrije Universiteit Brussel (VUB). The MOTUS platform is designed for the scientific collection of Study data by means of survey questions and diaries. MOTUS collects location information, both at the moment MOTUS is used (foreground), and when MOTUS is not used (background) by the respondent.

Scientific research is one of the various purposes for which the VUB processes personal data. Every scientific project within the VUB is subject to the VUB privacy policy [https://www.vub.be/en/privacy-statement]. MOTUS is explained as a specific project in this document. No reading of the VUB privacy policy is necessary to get the full information.

The Privacy Policy provides information to participants in MOTUS studies and any other interested participants about how their Personal data, Study data and User data are stored, preserved and processed. Respondents participate through their smartphone, tablet or

computer/PC/laptop, using the MOTUS web or mobile application to do so. In addition, the MOTUS applications can be used to capture data through sensors, or via an API-link with external databases (all collectively, the "Services"). MOTUS has no API-links providing data to external databases.

When reference is made to MOTUS later in this Privacy Policy, this means the MOTUS platform and its applications, and so the Services.

The Services are used for different studies. This document discusses the Privacy Policy that applies to all of these studies. As a supplement, more information can be provided in the Services for each study.

This policy was written in Dutch. To the extent a translated version conflicts with the English version, the Dutch version controls. Unless indicated otherwise, this Privacy Policy does not apply to third party products or services or the practices of companies that we do not own or control, including other companies you might interact with on or through the Services.

Data controller

The VUB, represented by the rector, is the controller for the processing of the Personal data of respondents. The VUB has appointed a Data Protection Officer who assists the VUB in monitoring all processing of Personal data that takes place at the VUB.

MOTUS and the Services are managed by the Research Group TOR of the VUB (VUB-TOR), led by Prof. Ignace Glorieux (Ignace.Glorieux[at]vub.be; +32 2 614 81 55). The project manager is Joeri Minnen (Joeri.Minnen[at]vub.be; +32 2 614 81 49). Specific contact details may be provided for each study.

Hbits CV acts as joint data controller. Hbits CV is a spin-off of the Research Group TOR and the Vrije Universiteit Brussel through the licence to MOTUS. The joint role as data controllers held by the VUB and hbits is set out in a contract. The MOTUS Privacy Policy therefore applies to hbits CV.

Scope and acceptance of this privacy statement

The MOTUS Privacy Policy applies to all data that could be used to identify an individual. When you use the Services (MOTUS web, mobile or other connected devices) and participate in the MOTUS studies, or when you provide your Personal data to MOTUS, you are deemed to accept the contents of this Privacy Policy.

If you do not accept this Privacy Policy, MOTUS and its applications may not be used and personal data may not be provided to MOTUS.

Collection of data

MOTUS makes it possible to collect, manage, store and process data from respondents. MOTUS receives information in a few different ways, including when you answer questions, register, track, complete or submit activities using the Services. Because the Services are used for various studies, large quantities of data are collected and managed. This may vary for each study. Only the information that is necessary in relation to the research questions is requested.

Personal data are collected, managed, stored and processed for one of the following purposes:

- Organising registrations for, invitations to and participation in studies;
- Drawing up group profiles;
- Conducting studies inside and outside the university and for statistical purposes;
- Contacting the winners of competitions or paying compensation in connection with a study;

- Verifying whether confidential material provided to MOTUS respondents is being used properly.

A Privacy Policy is designed to protect personal data. Because openness is important and because personal information can also be created by combining data sources, all data are discussed.

In what follows, a distinction is made between Personal data, Study data, Sensor data and User data that MOTUS collects, manages, stores and processes with the respondent's consent.

Personal data

Much of the Personal data is used to identify respondents as individuals, to contact them in the context of studies conducted via MOTUS and to secure and provide access to our Services. These data are first name, last name, sex, date of birth, language, address (street, number, postal code, city and country), email address, telephone details (landline, mobile phone) and the push tokens of the respondent's devices. Part can also be used as Profile information.

To make it possible to recall the details of conversations at a later time, in the context of a study communication with the respondent is also recorded.

Study data

The overarching goal of MOTUS is to capture study data on the behaviour, opinions, norms and values of respondents. This is done through having respondents independently answer questions and keep a diary, activating sensors (see further) and connecting relationships (e.g. family relationships). People's behaviour in daily life stands central. Respondents give consent to the collection of these data.

Upon consent data can also become available via data linking. Databases may be other collected databases within MOTUS, but also administrative databases or databases that provide more context about respondents' behaviour.

Research is only conducted on a combined dataset between pseudonymised or anonymised personal data, study data and user data such that it is not possible to identify the respondent. To use a complex term, this is called differentiated privacy. To give an even more complex explanation, differential privacy which is a system for publicly sharing information about a dataset by describing the patterns of groups within the dataset whilst withholding information about individuals in it.

Sensor data

Location information

Location information is being captured by our Services through the use of sensors available in the mobile application. To assess quality information on the daily life of people an accurate display of time and location is necessary. We collect and process location information when you sign up for and use the Services. Upon giving consent location tracking is active while you use MOTUS (foreground tracking) and also when you are not using MOTUS actively (background location tracking).

Foreground tracking is done to assess location information at the moment answers to questions are given, and when respondents register activities in their diary.

Background tracking and the processing of location information is used to build up a tentative data overview showing time, place and (if available) transport mode as necessary information in order to successfully complete a diary.

A diary length can be of various durations (a day, a few days, a week or even longer). During this period the MOTUS application might not always be used by the respondent (foreground tracking). Therefore background tracking is essential in providing a tentative data overview at the moment the respondent uses MOTUS. The tentative information is shown to the respondent and upon acceptance it becomes committed data. In the process of committing data tentative data can be corrected or supplemented. The committed data is used as Study data. The respondent stays in control of his/her own data, as the tentative data provides the respondent the necessary assistance to commit data to the dairy. The more accurate diaries are complete the higher the quality and the better personal feedback can be given back to the respondent.

The location data can also be used in interaction with a geofence. At the moment a geofence is entered, dwelled or exited the respondent can receive extra questions in relation to the specific location or region.

To protect the respondents privacy, location information is only stored during the phase when the respondent is asked to complete a task, during the period of completing a task or when interacting with a geofence.

User data

User data are collected during respondents' interactions with MOTUS. These User data are collected to provide a better service to respondents, but also to make it easier to contact respondents. As MOTUS is used, information is collected about when respondents use MOTUS and what devices and operating systems are used. This allows in-app messages to be sent.

Cookies

Cookies are used on the MOTUS website. Cookies are small data or text files that are placed on a computer by the websites that people visit. They are widely used to make websites work more efficiently and to provide information to website administrators.

MOTUS only uses first-party cookies. These are cookies that are created and placed by employees of MOTUS. MOTUS does not use third-party cookies. No information is forwarded to others via cookies.

The first-party cookies are:

- Session cookies: these are temporary cookies that are deleted when a browser is closed. In the case of MOTUS, they coincide with the necessary cookies, and are there to allow respondents to browse the MOTUS website and make use of its features.
- Permanent cookies: these cookies remain available to a browser until they are deleted. They are there in order to recognise returning respondents and make it possible to display information that respondents have previously provided. When a cookie is able to identify a respondent as a returning visitor, information such as login data does not constantly need to be re-entered.
- Necessary cookies: these cookies are strictly necessary for browsing the MOTUS website and making full use of its features.
- Cookies that send information to MOTUS: these cookies store certain aspects of respondents' browsing behaviour on the MOTUS website. The information that is collected concerns when and for how long a visitor uses the application, what

devices the user uses to log in and register data, and what browser or browser version is used. This information is used for support and statistical purposes.

Due to the composition of the cookies and the fact that only first-party cookies are used by MOTUS, no acceptance of underlying cookies by the respondent is required.

Meta data

To be able to analyse the data better, the data are also enriched with labels and further metadata information.

Transaction data

Transaction data are required in relation to the compensation that respondents may receive for their participation in studies. When a respondent is eligible to withdraw his or her compensation, a transaction takes place between MOTUS and the respondent. The transaction summary is retained for legal reasons and in order to demonstrate that these data and transactions are the subject of proper administration. We do not store any account information.

Children's personal data

MOTUS does not knowingly collect or solicit personal data from children under the age of 16. If MOTUS has inadvertently collected personal data from a child under the age of 16, it will remove such personal data from its archives as soon as reasonably possible. It is possible to collect personal data from children under the age of 16 directly from the parent or guardian and hence with his/her explicit consent.

Purpose of data collection

MOTUS is a research tool used for collecting study data with the aim of gaining more insight into people's behaviour and, from the central perspective of behaviour, gaining insights into opinions, norms and values.

Research questions may originate from educational institutions, government agencies, international authorities, non-governmental organisations, non-profit organisations, public corporations and commercial companies. They may be multidisciplinary, with relevance to multiple scientific fields.

Research is only conducted on a combined dataset between pseudonymised or anonymised personal data and study data such that it is not possible to identify the respondent.

Data storage

MOTUS uses Virtual Private Servers (VPSs) located in the EU that are only accessible to the project managers and the head of the IT department. The servers have been completely adapted by MOTUS employees and the database structures have been developed by MOTUS employees. As a result, MOTUS itself determines the types of data that come in and the security level.

MOTUS has maximum protection against theft, and even in the event of theft it would be almost impossible to retrieve and combine personal data, study data or user data since the information is spread across different servers. Digital keys are required to combine data. In order for data to be combined, a theft would have to take place simultaneously on different servers and in different locations, and even then, it would be difficult to assemble the data without prior knowledge. The same applies to data protection in terms of confidentiality (including against loss or negligence), but also in terms of integrity, accessibility and authenticity.

However, there is no such thing as a cast-iron guarantee. Should a theft, DDoS attack, loss, crash, or incident involving downtime or negligence nevertheless occur in spite of every effort, compensation may not be claimed for this, unless there has been negligence in meeting the obligations set out in the General Data Protection Regulation or GDPR. However, MOTUS warrants that data protection and security standards are continuously optimised to ensure an environment that is as efficient as possible at all times. The privacy of MOTUS respondents is the primary concern.

Data recipients

The VUB will not rent, sell or pass on personal data to third parties for financial gain. Data that can be traced back to individuals will only be provided to third parties if there is a legal basis or obligation for this, if the processing is necessary in order to represent the legitimate interests of the data controller or a third party, if it is necessary for the implementation of an agreement with the data subjects, or if the data subject has explicitly consented to the transfer of the data.

In accordance with the objectives of MOTUS, the personal data, study data and user data that are present within MOTUS are only available to the VUB, the Research Group TOR and the joint data controller hbits CV, a spin-off of the Research Group TOR.

The basis for data delivery is provision of study data. Data deliveries usually represent a limited part of the available study data related to the research question under consideration, and are supplemented with personal data or data derived from personal data.

The data may be made available in the exclusive or combined form of:

- Personal data
- User data
- Pseudonymised data
- Anonymised data
- Study data or a subset thereof
- Results

The VUB and hbits CV have access to all data and may combine them: personal data, study data and user data. Educational institutions and governments may submit a request for a subset of study data combined with pseudonymised or anonymised personal data, depending on the research question. NGOs, non-profits and public corporations may make similar requests, but only in combination with anonymised data. Commercial companies may only submit requests for results or statistical data, which in turn are available to everyone when they have been made public.

Personal data that are anonymised are recoded (reconstituted) in such a way that an individual cannot be identified. If an (anonymous) digital key is added so that the personal data can be made transparent again at a later time, this is referred to as pseudonymisation. Pseudonymisation would be possible, for example, if the client also needs future databases for its research question, and so wants to monitor changes in a personal profile and the combined study data.

It is important to stress that the respondent's identity is also preserved with pseudonymisation. The database user cannot find out the respondent's identity by means of pseudonymised personal data. The database can only be further enriched with additional study data and updates of the pseudonymised personal data with the cooperation of the MOTUS project manager. If pseudonymisation will be used in the study, the respondents will be informed and must give their consent.

Data retention period

Respondents are asked to participate in studies through MOTUS. After data collection has been completed and the databases prepared for scientific research, the personal data, study data and user data are anonymised in accordance with the law on the protection of personal data. Study data have a maximum retention period of ten years, unless stated otherwise when information is provided about the study and with the approval of the VUB's DPO.

If, in the context of a study, respondents need to be surveyed again at a later date, the personal data, study data and user data are pseudonymised. Respondents' personal data are stored within MOTUS. When respondents wish to be part of BEHAVE, they agree to the transfer of personal data, study data (composite or otherwise) and user data to BEHAVE. At that point, such respondents become BEHAVE panel members. As long as respondents are members of the BEHAVE panel, the data that BEHAVE collects, manages, stores and processes will remain available to BEHAVE for an indefinite period. When a panel member leaves, the collected data are also anonymised in accordance with the law on the protection of personal data. Again, a retention period of ten years applies to this anonymous data, unless otherwise determined in consultation with the VUB's DPO.

Rights with regards to personal data

Every individual, and hence every MOTUS respondent, has certain rights under GDPR.

The General Data Protection Regulation (GDPR), a European regulation that gives individuals far-reaching rights with regard to personal data and data protection, has been in force since 25 May 2018. In certain cases, and depending on the legal basis on which the VUB processes the personal data in question, individuals may be entitled:
- to ask what personal data about them are being kept and, if these data have been provided to the VUB by a third party, to inquire about the source of the data;
- to ask for data be rectified if it is incorrect;
- to object to the processing of their data;
- to request the transfer of their data to third parties or obtain a copy of them in a readable or machine-readable form;
- to be informed of the existence of any automated decision-making, and, if profiles are drawn up on the basis of this, to be informed of the underlying principles, importance and consequences thereof;
- if a number of conditions are met, to ask to be 'forgotten' by the organisation that has processed their personal data.

Further questions

If you have any questions after reading the MOTUS privacy policy, or if you have questions about your rights as a MOTUS respondent or how data are stored and processed, or in order to exercise your rights in relation to your personal data, contact the MOTUS project managers.

Compliance with privacy legislation and complaints

At a national level, the Data Protection Authority (DPA) is the authority that supervises compliance with all privacy legislation and to which complaints can be submitted about the processing of personal data. The DPA can be contacted at: Drukpersstraat 35, 1000 Brussels, or by telephone on +32 2 274 48 00, or by email at contact[at]adp-gba.be. More information about the DPA is available at www.gegevensbeschermingsautoriteit.be. An online form for complaints is available at: https://www.gegevensbeschermingsautoriteit.be/burger/acties/klacht-indienen

Information can also be obtained from and complaints submitted to the Data Protection Officer of the VUB, who can be contacted at dpo[at]vub.be.

17.3 Data Protection Plan of hbits CV

A Data Protection Policy (DPP) is designed to consider potential data protection and privacy impacts of processes and projects. It combines instruments like Data Protection Impact Assessment (DPIA), Data Management Plan (DMP) and Privacy Policy documentation. A Data Protection Policy is a deep dive to achieve protection goals but also a method for recording (identified) risks, issues and the safeguarding factors introduced to alleviate such concerns or potential impacts.

17.3.1   hbits CV

This DPP is related to the work of *hbits* (hbits CV). As part of its social responsibility *hbits* wants to operate in line with the data protection laws, regulation and rules as described in the General Data Protection Regulation (GDPR) and the national law(s).

hbits is a research company based in Belgium with address Witte Patersstraat 4, 1040 Etterbeek and company registration number (VAT) BE0712.734.818, and is related to the Vrije Universiteit Brussel (VUB) as the officially recognized Spin-Off of the Research Group TOR of the Sociology Department.

The main function of *hbits* is to collect insights on how people behave in their daily life and in which context their actions take place. Therefore *hbits* depends on the collection and analysis of information about people. The protection of peoples identity is one of the important actions to keep the respondents' and public's confidence.

To collect and analyze this information hbits mainly makes use of the MOTUS platform as it is able to collect survey data, time diary data and sensor data. *hbits* is the licensee of the MOTUS platform (more information https://www.motusresearch.io), and the related BEHAVE platform (more information https://www.be-have.io) which provides respondents for the researches ran via the MOTUS platform. VUB is the product owner of both platforms, and therefore besides *hbits* also the Research Group TOR of VUB uses the platforms for their research purposes. To settle this the usage of the platforms a Joint Data controller Agreement between VUB and *hbits* is in place.

For the projects and researches that *hbits* performs in relation to VUB projects and researches, the DPO of the VUB is in charge. In case of any data breach or complaints the VUB DPO is to be contacted (as well).

Nevertheless *hbits* holds a DPP as an independent research company. This document will describe the DPP of *hbits*, incorporating various other security and privacy related documents.

17.3.2   Introduction

Data protection laws require a DPP to be completed to assess proposed measures that pose particular risks relating to how personal data is used. This includes a (full) DPIA.

It is essential that each person, client, employee, supplier or other persons and/or organization that comes in contact with *hbits* has an oversight of the scope of the projects and researches that run within *hbits*. The responsibility for completing and respecting the DPP ultimately falls with the Executive Board of *hbits*.

This document describes the minimum standards of how personal data must be processed, collected, handled and stored to meet *hbits* data protection standards.

Anyone who works for *hbits* has a responsibility for ensuring personal data are collected, stored and handled appropriately.

Data users are obliged to comply with *hbits'* DPP when processing personal data on *hbits'* behalf. An agreement has to be signed defining possible disciplinary actions in case of a breach.

17.3.3   Identification of the need of a DPIA

Article 35 of the General Data Protection Regulation (GDPR) prescribes that a DPIA shall be conducted by a controller where a type of data processing, in particular using new technologies, is likely to result in a high risk to the rights and freedoms of individuals.

As a controller, under the GDPR an organization will need to assess, decide and document whether a DPIA is necessary for each proposed data processing operation. Although *hbits* states that no individual nor organization face a high risk to their rights and freedom it acknowledge that most of the projects or researches have a benefit in conducting a DPIA.

Especially when

- systematically monitoring, tracking and observing individuals' location and behavior, and
- combining, linking or cross-referencing separate datasets where such linking significantly contributes to or is used for profiling or behavioral analysis of individuals

are two of the elements on an EU-privacy regulators' list describing to have a high(er) risk to the rights and freedoms of individuals pursuant to GDPR Article 35(1).

17.3.4   Processing personal data

The main activity of *hbits* is to collect information on the behavior of people within a contextualized setting. During the respondents' participation personal data, research data and user data are being collect, stored and processed. All data are dealt with properly, irrespective of how they are collected, recorded and processed.

The Privacy Policy MOTUS gives a detailed overview on the personal data that are collected, as well as on the research data and the user data. This Privacy Policy is updated when necessary.

To arrive data security, *hbits* follows privacy- and security protection goals aligned with the data protection principles defined in article 5 GDPR. The goal is to receive and maintain the confidence of individuals. This document discusses seven protection goals within 2 level. Below the levels are explained.

GDPR considers respondents as 'Data Subjects'. However Data Subjects are a general term for respondents, employees/consultants, market contacts and organizations from which *hbits* holds personal data. This document will mainly discuss the protection of personal data of respondents (Hoorn & Montagner, 2018). They will also be addressed as 'individuals' and 'persons' and can be part of a group.

Risk of IT security

3 of the goals are related to IT security:

(1) Availability

is the requirement that personal data must be available and can be used properly in the intended process. The data must be accessible to authorized parties and the methods intended for their processing must be applied.

(2) Integrity

refers, (i) to the requirement that information technology processes and systems continuously comply with the specifications that have been determined for the execution of their intended functions. (ii) the data to be processed remain intact, complete, and up-to-date.

(3) Confidentiality

refers to the requirement that no person is allowed to access personal data without authorization. It ensures the protection against unauthorized and unlawful processing.

Data protection goals

4 of the goals are related to data protection goals:

(1) Data minimization

is the requirement to collect, process and use only personal data than are necessary for the achievement of the purpose of the processing.

(2) Unlinkability

refers to the requirement that data shall be processed and analyzed only for the purpose for which they were collected.

(3) Transparency

is necessary for the monitoring and control of data, processes, and systems from their origin to their erasure and is a prerequisite for lawful data processing. Transparency of the entire data processing operation and of the parties involved can help ensure that Data Subjects and supervisory authorities can identify deficiencies and, if necessary, demand appropriate procedural changes.

(4) Intervenability

refers to the requirement that Data Subjects are effectively granted their rights to notification, information, rectification, blocking and erasure at any time, and that the controller is obliged to implement the appropriate measures.

Below *hbits* shows how is dealt with these privacy- and security protection goals.

17.3.5   Data minimization

Automated processing operations following a data collection strategy

Via MOTUS *hbits* is particular focused on the collection of survey and time use data, and this from individuals or groups of people (like households, teams or organizations). Diary data combine activity data (What people do?) and contextual data (Where, With whom, How, Why, …).

Data can be collected actively and passively. Active means that data are registered by the respondent him/herself. Respondents make use of the MOTUS web and/or mobile application. Passive means that data is collected via internet connected devices or sensors,

or that external databases are linked. An example is location tracking from respondents. Active and passive data streams can be combined within a research and the gathered information can be presented to the respondents in the MOTUS applications.

In the MOTUS back-office the researches are developed: the research components (questionnaires, diary, context, communication) are defined, a fieldwork flow is designed for automated operations and data management is programmed and organized. MOTUS makes it possible to define multiple researches, that can run at the same time, even within the same respondent (for panel research purposes).

MOTUS is able to perform an important number of tasks and subtasks within the process of statistical production. To inform interested parties MOTUS is mapped on the GSBPM-architecture, developed by UNECE, EUROSTAT and OECD. The Generic Statistical Business Process Model (GSBPM) is a means to describe statistics production in a general and process-oriented way. It is used both within and between statistical offices as a common basis for work with statistics production in different ways, such as quality, efficiency, standardization, and process-orientation.

The figure below shows the main phases Build, Collect and Process in which MOTUS provides a service. Also in Design, Analyze and Disseminate MOTUS can play a role.

Figure 68: Main GSBPM-phases in which MOTUS provides support
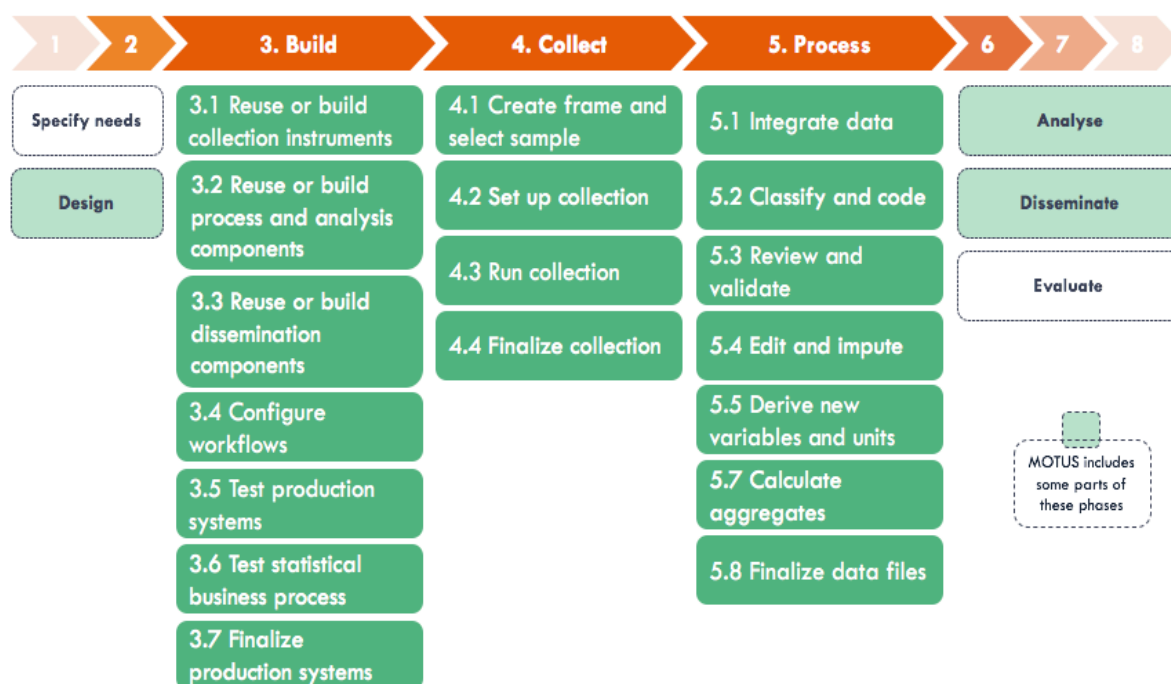
MOTUS collects new data, and can reuse earlier collected data. Respondents can be newly invited, but can also accept to be reinvited and even become a BEHAVE panel member. These respondents give consent for data collection and research purposes to store and use their personal, research and user data throughout their membership.

The amount of respondents that (can) interact with MOTUS (over time) is undefined.

Reduction of collected attributes of the data subject

hbits uses MOTUS in various projects. Each project can have its own specifics but the way in which data are collected and how the information flows run, does not exceed the information provided within this DPP. As an example not all researches make use of sensor data. In that case no passive data flow is necessary.

Each individual project has a detailed information fiche to inform the respondents. This information is presented during the invitation, on our webpage and in the applications. The data that is collected from the respondent is related to the research question within a research.

No more information is asked from the respondent then necessary, and the respondent has always an overview of the research tasks to be fulfilled within the research. All actions within the participation to a research are on a voluntary basis. At any time during the participation to a research respondents have an overview of the given information. Also, for each specific research a Privacy Policy (or an addendum to the General Privacy Policy MOTUS) is defined.

Respondents can contact the responsible researcher via a contact page, via email and when available via telephone.

Procedures for pseudonymization and anonymization

The effectiveness (and legality) of both anonymization and pseudonymization hinge on their abilities to protect respondent from re-identification.

To achieve this MOTUS works with Universal Unique Identifiers (UUID), and this on 2 levels. Every respondent in MOTUS receives an UUID, every research has a research UUID. Personal and research data are not present in one database. Only based on the UUIDs for respondents and researches personal data and research data can be merged.

When the data collection and data cleaning period is finalized personal data and research data are removed, unless different-wise determined in a contract, from MOTUS and stored in a data archive.

Storage period

A difference is made between personal data and research data. Research data can also contain anonymized personal data.

For the projects hbits executes independently of VUB, the storage period for projects and researches is defined as follows:

- Personal data
  Based on necessity (data collection, data cleaning), based on a contractual agreement with the respondent (e.g. BEHAVE panel)
- Research data
  Unlimited

For the projects that are executed together with VUB

- Personal data
  Based on necessity (data collection, data cleaning), based on a contractual agreement with the respondent (e.g. BEHAVE panel)
- Research data
  10 years

17.3.6   Transparency

Architecture and data flows

For the data collection and storing of the data MOTUS employs two types of architectures: a monolithic and microservice architecture. When a project or research is finished the data is transferred to a data archive.

Monolithic architecture

The first architecture combines in total 6 components or modules:

- Backend server: the backend server stands central in the MOTUS platform. It holds the database, the back-office API and the client API.
- Back-office: the back-office serves as the research environment where the researcher sets up a research and the fieldwork can be followed. The back-office runs in a browser.
- Analyze server: the analyze server holds a replicate of the database of the backend server and prepares the reports for the backend server, which at its part can be called by the back-office.
- Back-up server: the back-up server is a replicate for secure storing from the backend server and the analyze server.
- Client portal: the client portal holds the MOTUS-web application and an underlying webserver.
- Mobile devices: the mobile application is available for Android and iOS.

All together this is called the MOTUS platform.

Figure 69: The MOTUS platform
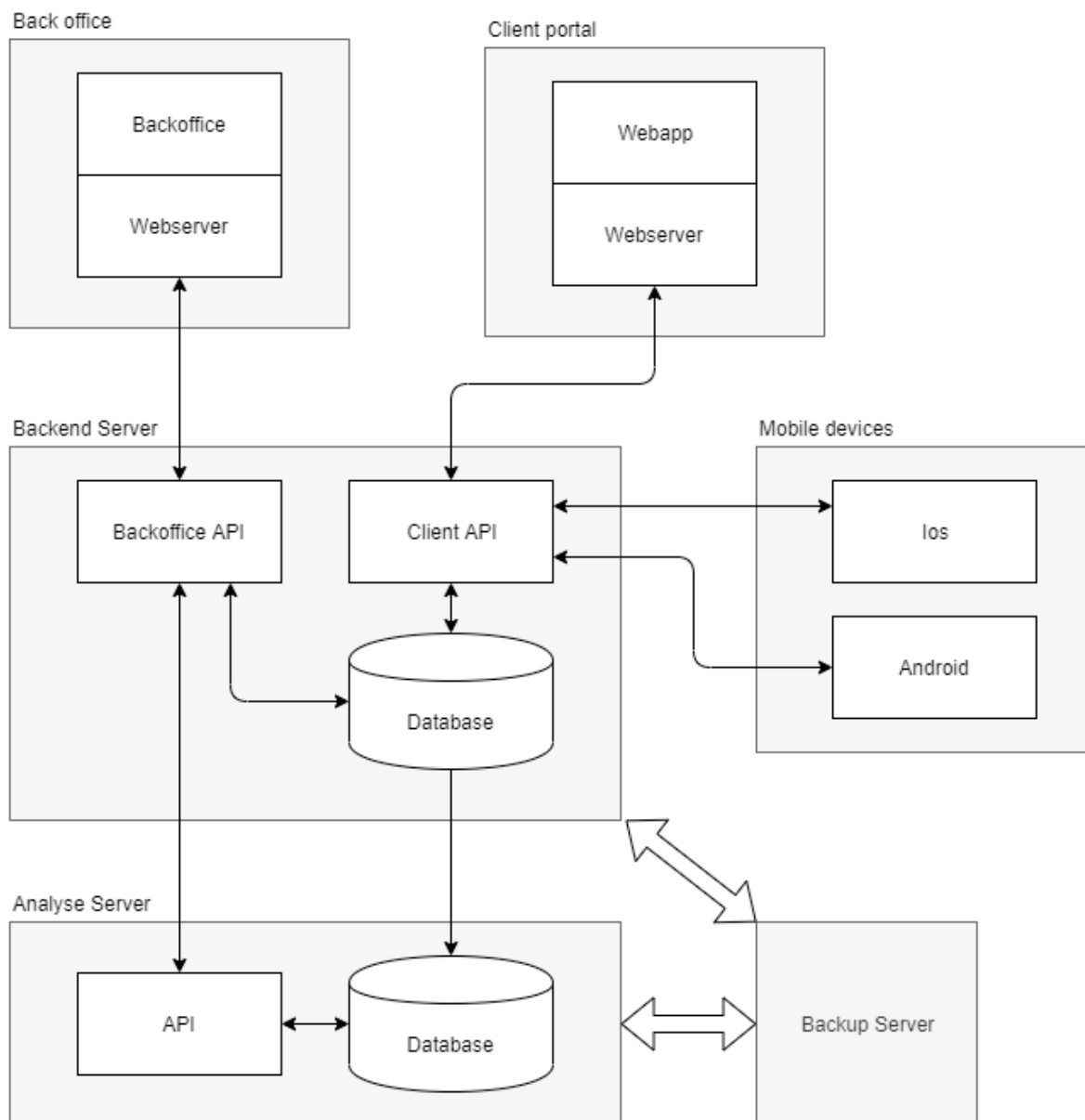
There are three API's that arrange the entrance to the components:

- Backoffice API: both ways webserver back-office and analyze server
- Analyze server API: both ways database (to prepare reports) and back-office API to send over reports and other analytics.
- Client API: Receives the input from the web & mobile app and syncs the data on both applications. It could also function as a data harmonization tool.

Microservice architecture

Microservices are small and independent services handling a particular problem or performing a certain/specific task. Via an API the collected data is presented to the monolithic architecture.

These microservices can be external data sources to MOTUS. However often these microservices are developed in function of MOTUS.

Figure 70: MOTUS Microservice architecture
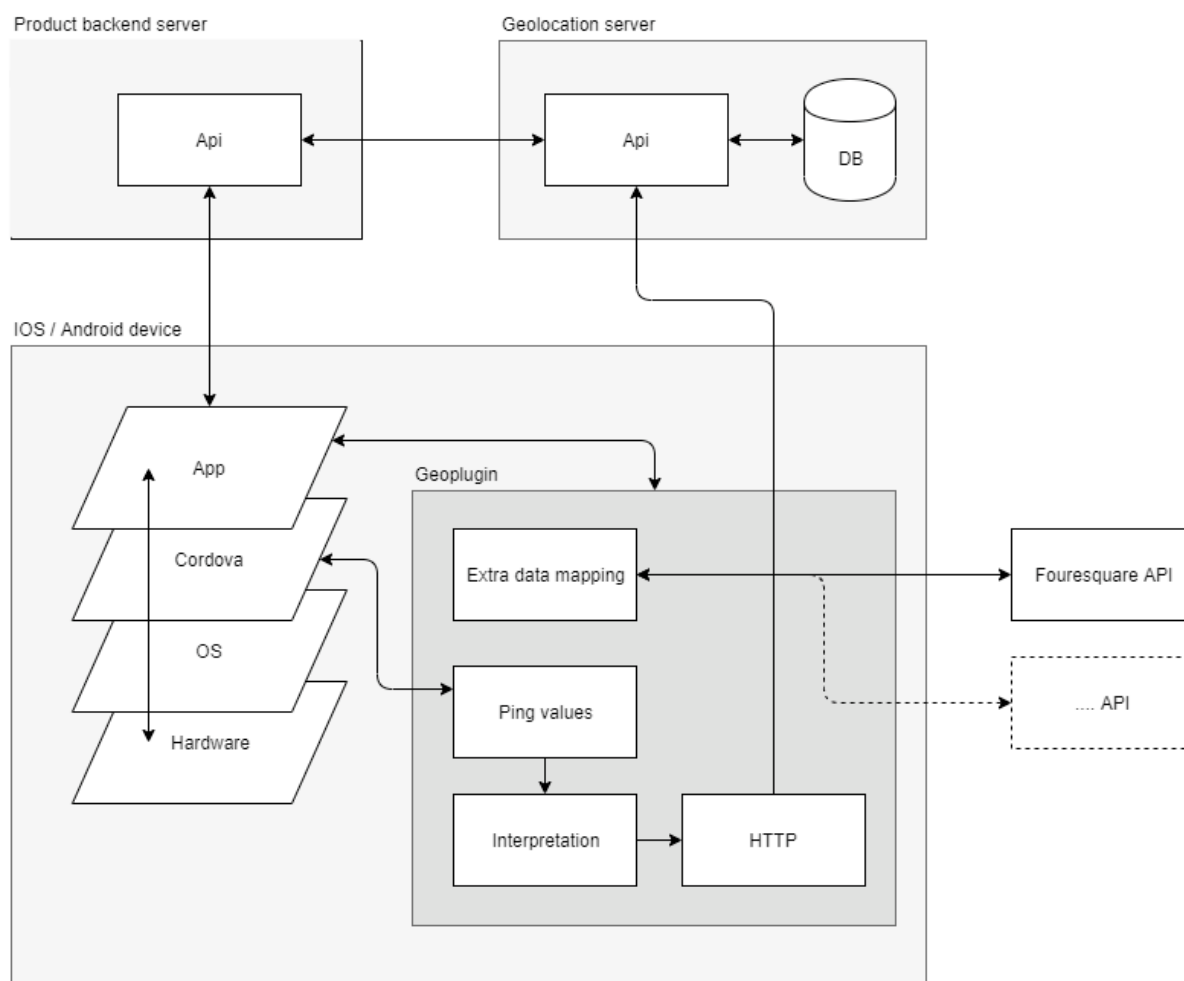
An example is location tracking where the sensor of a respondents' smartphone collects information on the position of the respondent. Also external databases can be connected, like Foursquare or Google to provide extra context.

Every microservice has its own server, database and processing. The connection with the MOTUS core is done via a Rest API. In this way no (personal) data is communicated with the microservice, only the UUID to be able to match the data in the backend server of the monolithic architecture.

Data input and output

MOTUS is a protected environment. All input and output is controlled. Nevertheless there are 3 processes that are linked to MOTUS :

(1) provided input to MOTUS (eg. registration, opt-in, administrative data, sensor data, …)
(2) output internally used within MOTUS (eg. invitation emails, internal dashboard, …)
(3) output made available by MOTUS to externals (eg. research data, external dashboard, …)

Data archive

After the data collection and cleaning phase personal data and research data are erased from the MOTUS environment.

Data are stored in a self-hosted data archive.

Privacy by design and default

hbits uses a Privacy by Design and Default approach in all its work, but in particular when:

- building new IT systems for storing or accessing personal data;
- developing new applications or research approaches;
- embarking on a data sharing initiative; or
- using data for new purposes.

Privacy by design is an approach to projects that promotes privacy and data protection compliance from the start. It is a key consideration in the early stages of any project, and then throughout its lifecycle.

Taking a privacy by design approach is an essential tool in minimizing privacy risks and building trust and will designing projects, processes, products or systems with privacy in mind from the outset .

External service providers and third parties

Server provider

hbits makes use of the services of Contabo. MOTUS runs on VPS certified servers.

More information can be found on the service's website: https://www.contabo.com

Email service

hbits makes use of the services of Mailgun to send and track emails to/from respondents.

More information can be found on the service's website: https://www.mailgun.com

External databases

There are various external databases *hbits* works with:

- Administrative databases provided to hbits

    These databases can contain personal data, research data or user data. The acceptance of this data is based on the given consent to the provider of these data.

    For each research the respondent is notified from which organization the data is received in order for the respondent to make a claim on their respective rights. An example is employee data from the VUB.

- Content databases

These databases contain extra context to the answers/information the respondents have provided to MOTUS.

For each research the respondent is shown an overview of the services that are used within that research, and how it relates to the protection of their data. An example of a content database is the Places API of Google, or the OverPass API of OpenStreetMap.

MOTUS acts as an in-between the respondent and the external service. MOTUS anonymizes the respondent in contact with the external service provider.

Data transfer contracts

hbits is a spin-off of the Research Group TOR of VUB. Via a joint controller agreement Personal Data is shared with and can be processed by VUB. This makes VUB an important stakeholder. VUB and hbits can also outsource work between the two organizations.

Other important stakeholders are statistical organizations (like Statbel or Eurostat), which are important due to a joint development trajectory.

In case of a transfer of personal data recipient can only use the data for the defined purposes. The receiving party must maintain a data protection level in line with the GDPR to be able to accept the data.

Where personal data are transmitted by a third party (like a sample supplier) to *hbits*, these personal data can only be used for the intended purpose.

There are no other contracts or agreements for the transferring of personal data. Personal data can be merged with research data only when pseudonymized or anonymized.

The conditions for transferring research data are specified in the Privacy Policy MOTUS. *hbits* acts conform the EU data protection shield.

Overview of consents and objections

A lawful base for processing personal data is obligatory. *hbits* requests a lawful base for every research.

Each running research via MOTUS is listed, and the consents are provided via https://motusresearch.io. In case of objections a summary is provided.

Non-MOTUS research has an informed consent as well, but are not part of the list.

17.3.7   Availability

Data storage and backup

The following server settings apply:

For the backend server:

- VPS
  Linux: Debian 9 // Updated to CentOS (free version of Redhat)
- Ten core CPU
- 60GB RAM
- Disk storage: 1600GBPort/bandwidth: 1Gbit/s

For the backup server:

- VPS

- 2x100GB back-up space
- Internal network only

Repair strategies and alternative processes

Data collected via the backend VPS server of MOTUS is automatically replicated at the backup VPS server.

Pseudonymized or anonymized data is stored in a data archive in relation to the storage time. Archived data are accompanied with meta-data information and a technical report. A technical report includes also a description of privacy preserving actions.

Configurations and data structures

MOTUS uses a relational database to store all gathered information of the respondents. Personal data and research data are not present in the same database. The matching of the information can only be done via source code and MOTUS internal processes, or manually via an UUID-key. This key is only available to the Admin role.

MOTUS data is stored in a MySQL MariaDB, running on a certified VPS-servers. The performance, updates and security level of the servers is monitored via IT specialists working for *hbits*.

Protection against external influences

MOTUS takes into account a number of security measures.

Connection to the MOTUS backend database is only allowed from application or analysis server. No external connections to the database are allowed.

Each user application (respondent web frontend, web app, registration website, …) has a unique API key, with configurable rights.

Data transmission is done via a respondent unique UUID key. Data is transmitted over a https protocol which an SSL/TLS encryption layer. The SSL connection encrypts and decrypts requests and responses.

The JSON encoding standard is used to transmit pieces of data between the MOTUS-components.

Actions are undertaken against DDoS attacks.

17.3.8 Integrity

Maintainability strategy

Maintainability deals with the easiness to maintain and (even) further develop the software platform. One aspect is the choice of the programming language. For its front and back-office MOTUS makes use of the frameworks Angular, Ionic, jQuery and Koseven.

The code is documented in Github, and uses code history to track changes in the source code. Particularly important is that the logic between the web and mobile logic is stored in a common library to improve maintainability, and/or to organize the work between collaborators, whereas the design code has been separated for web and mobile.

Another aspect is the organization of responsibilities via a release management process. This process takes in to account the following elements:

- Deployment strategy

MOTUS today uses a 2-stage deployment strategy having a development and production environment. In the future a more phased deployment would have benefits including substages for rollout, testing, and rollback in case of problems.

▪ Modularity client-server

MOTUS aspires to a 3-tier application architecture. A 3-tier application architecture is a modular client-server architecture that consists of 3 parts:

- a presentation/client tier: includes the graphical user interface, and communicates (API) with the other tiers
- an application tier: includes the business logic
- a data tier: includes information storage

▪ Agile management

hbits organizes development stories via an iterative and incremental method of managing new stories, new actions and bugs using Github and Jira. *hbits* also applies (the development of) Unit Testing.

Documented assigned of rights and roles

Roles are assigned on the Admin level, the group level and the research level. These levels are documented in the MOTUS manual. On the research level another set of rights are defined, 12 at the moment.

On the research level rights and roles are based on the task profiles of an employee or researchers.

An overview is documented of the rights and roles permitted.

Workflow and profiles

An user guideline on the back-office of MOTUS is available to employees, both of *hbits* as for the Research Group TOR of VUB.

Research components are inventoried in a library and research flows are described in a technical and meta-data documentation.

hbits provides training to its employees and researchers of the Research Group TOR of VUB.

Security tests

MOTUS has been tested by T-Systems Multimedia Solutions (Dresden).

17.3.9 Confidentiality

Data secrecy

Personal data are subject to data secrecy. Any unauthorized collection, processing, or use of such data by employees/consultants is prohibited. Any data processing undertaken by an employee/consultant that he/she has not been authorized to carry out as part of his/her legitimate duties is unauthorized. The "need-to-know" principle applies. Employees/consultants may have access to personal data only as is appropriate for the type and scope of the task in question. This requires a careful breakdown and separation, as well as in limitation, of roles and responsibilities.

Employees/consultants are forbidden to use Personal Data for their own private or commercial purposes, to disclose them to unauthorized persons, or to make them available in any other way. Supervisors must inform the employees/consultants at the start of the employment relationship about the obligation to maintain data secrecy. This obligation shall remain in force even after employment has ended. The employment/consultant agreements contain appropriate confidentiality obligations.

Assignment of rights and roles

4 environments are of importance:

(1) Backoffice of MOTUS
(2) Repository of MOTUS
(3) Server and database
(4) Data archive *hbits*

MOTUS back-office

This environment is also role protected.

Admin role MOTUS:

- Joeri Minnen – Managing Director
- Ignace Glorieux – Professor at Research Group TOR
- Theun Pieter van Tienoven – Research Director at Research Group TOR

Admin role Research:

- One research responsible for each research

Role distinction between two teams responsible for:

- Data collection
- Data valorization

Besides the Admins no one can be part of both teams within the same research. For every person a more specific role distinction can be specified based on the task description of the person.

Repository source code

The source code is stored in a Github private repository. The Github is accessible for the IT-developers of MOTUS.

Server and database

The servers of MOTUS are present in the Contabo Data Center in Germany. The control panel is available for the IT-developers of MOTUS.

Unauthorized physical entrance is not possible.

Data archive

The data archive runs on a self-hosted server.

The entrance to the data archive is protected via credentials and a 2-step verification. Entrance is allowed for:

- Joeri Minnen – Managing Director
- Ignace Glorieux – Professor at Research Group TOR
- Theun Pieter van Tienoven – Research Director at Research Group TOR

Definition of organizational procedures

Every *hbits* employee/consultant has signed a confidentiality contract. The same counts for the researchers of the Research Group TOR of the Vrije Universiteit Brussel.

Implementation of a secure authentication process

Both the respondents and the researchers have to authenticate against either the front- or the back-office to make use of the service. In both instances confidentiality and data integrity are essential.

To meet this criteria, user credentials are issued by MOTUS while these settings can be modified for every different research. Initial credentials are provided to the user by letter or e-mail. Users can change their username/password via the applications. As a setting this password change can be mandatory.

Passwords are encrypted in the database. Each user application has his own unique API-key, as is the same for every respondent. Username and identifier are connected in the database. Connection to the database is strictly ruled.

MOTUS front-office:

- Login with username and/or email + password
- Logins are linked to researches
- Password are encrypted in the database
- Initial passwords are auto-generated and can contain numbers, characters and signs

The back-office of MOTUS for researchers requires an extra 2-factor authentication.

MOTUS back-office:

- Login for admins with username and/or email + password
- Password are encrypted in the database
- Requires 2-factor authentication
- User role restrictions

Backup tokens can also be created.

Encryption strategy for stored and transferred data

Passwords and recovery codes are stored by concatenating with the salt and encrypting the result. Information is not encrypted in the database to retain a high performance of the database in the communication with the different front-office options.

Data is transmitted over a https protocol which an SSL/TLS encryption layer. The SSL connection encrypts and decrypts requests and responses.

Any respondent can ask by a simple request to be deleted. If a respondent is deleted from the database all data that is linked to this respondent is deleted from the research and database. MOTUS does not apply soft deletion, the data is fully removed.

17.3.10 Unlinkability

Restriction of processing, utilization and transfer rights

MOTUS is a research platform helpful in the organization of different researches, in the build-up of a research, in the collection and the processing of data.

The qualities of the MOTUS platform can be used for every research. This leads to efficiency but also to better protocols and harmonization of input and output.

Transfer of research components via a library is freely available. The transfer of personal data and research data can be restricted. Besides roles and rights on the person level also rights can be defined on the research level and on the group level. Individual researches can be made invisible and/or can be locked. When researches are related to a group of users, all researches of the group are visible or locked to others.

If the entrance to researches are restricted, no personal or research data can be seen, downloaded or transferred.

Separation of organization/departmental boundaries

hbits is at this point in time a fairly small research company, and the Research Group TOR of VUB is group of researchers, pre-docs and post-docs. A clear separation between *hbits* and TOR is achieved. For the projects and researches to which *hbits* and TOR share responsibilities a joint controller agreement is in place.

The Admin is able to put further restrictions on the boundaries between the current work organization and work flows, and will update these restrictions/boundaries in the future when necessary.

Anonymization and pseudonymization strategy

At the moment respondents credentials are created (username and password) also a UUID is attached to the respondent. With this UUID different kinds of information stored and collected from the respondent can be linked during the data collection and data cleaning process.

After the data collection is finalized 2 strategies can be applied:

- Anonymization: information which does not relate to an identified or identifiable natural person or to personal data rendered anonymous in such a manner that the Data Subject is not or no longer identifiable. Also the UUID of the respondent is transformed.
- Pseudonymization: processing of personal data in such a manner that the personal data can no longer be attributed to a specific respondent without the use of additional information, provided that such additional information is kept separately and is subject to technical and organizational measures to ensure that the personal data are not attributed to an identified or identifiable natural person. Personal data are stored separately together with the UUID of the respondent. The new database holds the anonymized data and the UUID. Linkage between with the personal data remains possible on a later moment.

For each research the respondent is informed about the strategy and the consent of the respondent is asked.

17.3.11   Intervenability

Every respondent has the following rights. Their request is to be handled immediately by *hbits* and may not result in any disadvantage to the respondent. Where the relevant personal data are not being processed by *hbits* the relevant client contract must be consulted in respect of any process to be followed and the client has to be informed about such request immediately.

- Right of access

  The respondents may request information on which personal data relating to him/her have been stored, how the data were collected and for what purpose. If personal data are transmitted to 3rd parties, information must be given about the identity of the recipient or the categories of recipients.

- Right to rectification

  If personal data are incorrect or incomplete, the respondent can demand that they are corrected or supplemented.

- Right to withdraw consent

Where the personal data are processed on the basis of consent, the respondent can object to the processing at any time. These personal data must be blocked from the processing that has been objected to.

- Right to erasure

The respondent may request his or her data to be deleted if the processing of such data has no legal basis, or if the legal basis has ceased to apply. The same applies if the purpose behind the data processing has lapsed or ceased to be applicable for other reasons. Existing retention periods and conflicting interests meriting protection must be observed.

- Right to object

The respondent generally has a right to object to his/her data being processed and this must be taken into account if the protection of his/her interest takes precedence over the interests of the data controller owing to the particular personal situation. This does not apply, if a legal provision requires that the personal data are data to be processed.

- Right to data portability

The respondent has the right to request for the personal data provided by him/her to be made available to such respondent in a easily readable format, like a Word or Excel document.

17.3.12 Legal grounds for Data Processing

hbits collects, process and uses personal data only under a legal basis. A distinction is made between respondents, data provided by clients, employee data and marketing contacts.

Respondent data

Respondents are the most common Data Subjects for hbits. The legal grounds to collect process and use personal data are related to:

- Informed consent

Any freely given, specific, informed and unambiguous indication of the data subject's wishes by which he or she, by a statement or by a clear affirmative action, signifies agreement to the processing of personal data relating to the respondent.

The respondent gives a clear affirmative act by means of an active motion or declaration. Examples are ticking a box on a (digital) form, or a written or oral statement, which clearly indicate the respondents acceptance of the proposed processing of their personal data.

- Contractual relationship

Personal data can be processed within the context of a contract to which the respondents is a party, to fulfil relevant obligations and rights.

This is the case when respondents are willing to be part of the BEHAVE panel, which is connected to MOTUS. More information is available in the Privacy Policy BEHAVE.

hbits mainly uses a web app, mobile app and connected devices to collect personal data and to process and use this information. The respondent is informed by the Privacy Policy MOTUS and the information fiche for every specific research.

The privacy statement and any cookie information are available so that it is easy to identify, directly accessible, easily understandable and consistently available by and for the respondent.

The same counts for the tracking of respondents via the web app, mobile app and connected devices. *hbits* uses a pseudonym for the respondent, but nevertheless the respondent can opt out of this data collection.

Personal data provided by clients

hbits is considered the processor when personal data are provided by a client, and can only process these data in line with the instructions agreed with or received from the client.

The received data are subject to the same data protection goals that are applied within *hbits*.

Employment relationship

Personal data is processed to initiate, carry out and terminate the employment agreement. This also applies to consultancy agreements.

Marketing contacts

hbits treats marketing contacts no different than respondents in respect of the privacy protections accorded to them.

17.3.13 Data incident and recovery plan

Security breach

A security breach happens when personal data has been lost, or there has been unlawful processing (security incident). This happens when there is:

- A destruction or loss of personal data
  E.g.: Data center fire, accidental deletion of a file without backup
- Damage, unauthorized access, incorrect provision
  E.g.: Loss of USB stick, stolen laptop, malware, hack, email to wrong recipient

Breach management

The following actions apply to prevent incidents

- Automated check for hack, malware, …
- Reporting every internal incident to a fixed contact person
- Creation of internal procedures and training/awareness
- Agreements on reporting external security incidents

To this end a registration form and a key personal contact list is available. This also includes a list of external contacts if necessary.

Breach investigation

The following questions apply to get a full picture of the security breach:

- Description of what exactly has happened to the data
- Description of the nature of the personal data affected
  E.g.: Special data, personal data of sensitive nature
- Description of the extent of the incident
  E.g.: Number of people affected, amount of data affected per person, is the affected data shared within a chain?
- Description of the category of persons affected
  E.g. Respondents, employees, clients, market contacts

- Description on the impact on those affected
- E.g. Vulnerable groups, financial loss, material/immaterial damage

Breach repairing

Measures to limit the consequences of the incident

The flowing measures apply:

- Reporting to the DPO and the Authority for the Protection of Personal Data.
- The persons affected by the data breach must be informed if the security breach is likely to have an adverse effect on their privacy.
- Inform affected persons on how to protect themselves to this infringement.

In case of negative consequences for the affected persons a report must be done withing 72 hours after the security breach to the Authority for the Protection of Personal Data and the DPO.

Measures also involves communication to the outside world if necessary. It also involves the coverage for third-party liability, own costs, and possible fines.

Prevent similar incidents in the future

Prevention measures apply:

- Plan updating
- Plan documentation storage
- Improve backup strategy
- Renew risk management
- Document triggering events
- Plan emergency response team

17.4 Penetration test MOTUS

A penetration test or pen test has been performed to evaluate the security of the MOTUS data collection platform, and to identify weaknesses which if not solved could provide unauthorized parties access to the software platform and the database.

The penetration test was performed by T-Systems Multimedia Solutions (Dresden), and included two retest.

The document provides information on which initial vulnerabilities were found. At this point, all critical and high findings are solved, and the MOTUS data collection platform therefore passes the penetration test.

17.4.1   Objectives

The objects of MOTUS analyzed during the penetration tests were the:

- web app back-office
- web app front-office
- mobile app front-office

The web applications test coverage was:

- Cryptography (eg. check of transport encryption and the validity of the used certificate)
- Authentication Testing (eg. verification of secure authentication and registration mechanisms

- Authorization Testing (eg. examination of unauthorized access to protected areas; missing or insufficient authorization checks)
- Configuration and Deploy Management (eg. test network/infrastructure/platform configuration; review old, backup and unreferenced files for sensitive information; testing of supported HTTP methods; testing for presence of HTTP security headers, testing of outdated software)
- Data validation (eg. test of handling user input; testing for various injection possibilities, testing for user-controlled resource allocation – denial of service)
- Business logic testing (eg. check for weaknesses in business logic and processes)
- Client-side testing (eg. cross origin sharing)
- Error handling (eg. test for information disclosure)

The mobile applications test coverage was:

- Data storage and privacy (eg. sensitive data in local data storage; third party interactions)
- Cryptography (eg. key management use if deprecated cryptographic algorithm
- Authentication management (eg. 2-factor authentication; testing stateless authentication)
- Session management (eg. misues of rights, takeover of existing sessions)
- Network communication (eg. attacks based on communication weaknesses; man-in-the-middle attacks; verify TLS settings; certificate stores and SSL pinning)
- Code quality and build settings (eg. review of signing certificate information; existence of debugging code/mode)
- Anti-reverse engineering (eg. usage of simple obfuscation)

The method used in this document to determine the severity of vulnerabilities found is based on the CVSS approach. CVSS stands for Common Vulnerability Scoring System and is an industry-standard for calculating the severity of vulnerabilities (https://www.first.org/cvss/).

Severity levels are calculated based on identifying the characteristic of vulnerability that have a risk impact. These characteristics are further divided into two factors namely, probability of occurrence (Likelihood) and impact.

Only the CVSSv3 Base Score is used to determine the severity level. The numerical values (0.0 to 10.0) result in the following severity levels.

Table 13: CVSS Score and CVSSv3-Rating

CVSS Score	CVSSv3-Rating
0.0	Informational
0.1-3.9	Low
4.0-6.9	Medium
7.0-8.9	High
9.0-10.0	Critical

In addition, for each identified vulnerability, a subjective assessment of the tester ("tester rating") is given, based on experience. The assessments include characteristics of the vulnerability or environmental properties that are not considered in the standardization of the CVSSv3.

17.4.2   Test results table

Below the test results are given, at first and third occasion. When after the last test an open ticket was solved it got marked by Destatis as 'Approved'.

Table 14: CVSSv3-Rating: test results table web application

Description	Rating 1th test	Rating 3th test
Web application (frontend, backend)		
Command injection/Remote code execution – PHP Script	Critical	Closed
OS command injection (also known as shell injection) is a vulnerability that allows an attacker to exe- cute arbitrary operating system commands on the server to take complete control of an affected system with the privileges of the user running the application. Command injection attacks are possible when an application passes unsafe user supplied data (e.g. forms, cookies, HTTP headers etc.) to a system shell. These kinds of attacks are possible largely due to insufficient input validation. The ability to trigger arbitrary code execution over a network (especially via a wide-area network such as the Internet) is often referred to as remote code execution (RCE). Usually, this behavior is not intended by the developer of the application. A remote code execution also can lead to a full compromise of the vulnerable application, server or infrastructure.		
Stored cross-site scripting (XSS) – Pages	High	Closed
Stored cross-site scripting (XSS) – Email templates	High	Closed
Stored cross-site scripting (XSS) – Message template	High	Closed
Stored cross-site scripting (XSS) – Research information	High	Closed
Stored cross-site scripting (XSS) – Survey questions	High	Closed
Stored cross-site scripting (XSS) – System pages	High	Closed
Stored cross-site scripting (XSS) – System mails	High	Closed
Stored cross-site scripting (XSS) – API-keys	High	Closed
Stored cross-site scripting (XSS) – Maintenance	High	Closed
Stored cross-site scripting (XSS) – Settings	High	Closed
Stored cross-site scripting (XSS) – Merge fields	High	Closed
Reflected cross-site scripting (XSS) – CSV upload	High	Closed
The term "Cross-Site Scripting" (XSS) denotes the injection of malicious HTML or JavaScript code into a website. Such an attack is possible if a website receives data from non-trustworthy sources (e.g. URL or POST data) and reflects them directly or indirectly (after storing them to database) to a web browser without prior validation or proper encoding. Thus, XSS allows an attacker to inject and execute code in the browser of his victim. An attacker can use injected HTML or JavaScript code to steal session data and hijack a user session, he can change the layout or behaviour of the application or redirect the user to a malicious page. Furthermore, Cross-Site Scripting can be used to exploit browser vulnerabilities.		
Cross-site request forgery (CSRF)	High	Approved
Cross-Site Request-Forgery is an attack whereby a valid request to the server is being invoked by a remote attacker in the victim's browser session. A precondition for this attack is that the victim must be logged in the web application. If a victim can be forced to visit an attacker-controlled page which contains, for example, a JavaScript code snippet with a POST request on the vulnerable page, the action will be executed with the session that the user has on the vulnerable web application.		
Cacheable responses	Medium	Approved
Unless directed otherwise, browsers may store a local cached copy of content received from web servers. Some browsers, including Internet Explorer, cache content accessed via HTTPS. If		

	sensitive information in application responses is stored in the local cache, then this may be retrieved by other users who have access to the same computer at a future time.		
Missing cookie security flags		Medium	Approved
	When a web application sends a cookie to the client's browser, it may set several flags in addition, which enhance the security of the cookie and protect it from being stolen or manipulated. If the HttpOnly flag is included in the HTTP response header, the cookie cannot be accessed through client-side script. As a result, even if a Cross-Site-Scripting (XSS) flaw exists, access to the cookie is not possible through JavaScript. If the Secure flag is set on a cookie, then browsers will not submit the cookie in any request that uses an unencrypted HTTP connection, thereby preventing the cookie from being trivially intercepted by an attacker monitoring network traffic. The use of a cookie to the same-site context could be restricted by setting the 'sameSite' cookie attribute. The attribute can be set to two values: lax and strict. Setting sameSite=strict for a cookie implies that it can only be used in the context of the same domain. This means that the cookie is only sent if the requested domain matches the domain currently displayed in the URL bar of the user's browser. A cookie with the sameSite=strict attribute set is never sent with a cross-origin request. The value lax, in contrast, allows some exceptions regarding the sending of cookies via cross-site requests. By setting sameSite=lax for a cookie, the browser sends this cookie even if there is an action that changes the URL in the address bar of the user's browser, so-called top-level navigations (e.g. links), and one of the following http methods is used to send this request: GET, HEAD, OPTIONS or TRACE.		
Information disclosure – stack traces		Medium	Closed
	Information like software version numbers, software in use, names of developers or staff, stack traces, file or directory listings, etc. may help adversaries to collect and assemble enough information to enable them to push their attacks further against a system, application, an engineering or staff team.		
Weak password policy		Medium	Approved
	Applications should encourage the use of strong passwords and passphrases. An application must only accept passwords, which meet the following requirements: • Minimum length of 8 characters • Consist of characters from at least three of the following character sets: lower and uppercase letters, digits and special symbols An application must verify, that new passwords meet the requirements above or otherwise reject it.		
Information disclosure – HTTP responses		Informational	Closed
	Information like software version numbers, software in use, names of developers or staff, stack traces, file or directory listings, etc. may help adversaries to collect and assemble enough information to enable them to push their attacks further against a system, application, an engineering or staff team. Such information may be contained inside source code comments, HTTP headers, error messages, on status sites, log files, downloadable office document formats (PDF, DOCX, etc.) and so on. Every bit of information will help adversaries to prepare more sophisticated attacks against their targets.		

Session fixation	Medium	Approved
Authenticating a user, or otherwise establishing a new user session, without invalidating any existing session identifier gives an attacker the opportunity to hijack authenticated sessions.		
Weak password management – Missing brute-force protection	Medium	Approved
Brute-force and dictionary attacks are used to ascertain valid username / password combinations, which then can be used to log into an application or service. These attacks are based on automated guessing by simply checking all possible combinations of a selected character set (brute-force) or pre-defined list (dictionary). If an application does not implement sufficient measures to prevent multiple failed authentication at- tempts within a short time frame, it is more susceptible to brute-force attacks.		
Missing session timeout	Low	Approved
Sessions are used to cache the authentication of a user to a system or application. The application must automatically log out a user if that user has been idle for a certain amount of time, ensuring that it is not possible to "reuse" the same session and that no sensitive data remains stored in the browser cache.		
Privacy violation: shoulder surfing	Low	Approved
A password need not be viewable to its owner and must not be viewable to others. If a password is displayed in plain text, anyone in the vicinity could see and use it to compromise the system. In computer security, shoulder surfing refers to using direct observation techniques, such as looking over someone's shoulder, to get information. Shoulder surfing is particularly effective in crowded, public environments. This threat particularly applies to mobile devices, which are generally intended for use in all environments, both private and public.		
Outdated libraries	Medium	Closed
Security vulnerabilities are detected throughout the life cycle of any hard- or software product. If the product is still supported, vendors usually provide updated versions which fix known vulnerabilities. Failing to update installed software components allows adversaries to exploit these vulnerabilities, possibly providing access to sensitive data or compromising the application. The application includes outdated JavaScript libraries with publicly known vulnerabilities. Further- more, the usage of old software indicates that no process for active patch management is implemented.		
Denial of service	Low	Approved
A Denial of Service (DoS) attack tries to overload a system or network with the goal to limit the availability to its regular users and thereby inflicting damage because the system is not available. Targeted resources are a service's CPU, volatile memory (RAM), persistent memory (hard disk), network band-width or socket count. A DoS attack may inflict temporal disruption (by the end of the attack, the service returns to its normal state) or persistent disruption (services are stuck in an overload state, even after the attack has ended; accounts may be locked, storage be depleted).		

Table 15: CVSSv3-Rating: test results table mobile application

Description	Rating 1th test	Rating 3th test
Mobile application (iOS, Android, mobile API)		
Missing root/jailbreak detection	Medium	Medium
Rooting (Android) or Jailbreaking (iOS) is the process of allowing users of a mobile device to attain privileged control (known as root access) over various subsystems. Rooting a device gives similar access to administrative (superuser) permissions as on Linux or any other Unix-like operating system. Like most other defences, root detection is not very effective by itself, but implementing multiple root checks that are scattered throughout the App can improve the effectiveness of the overall anti-tampering scheme. From a security perspective, rooted devices are a particularly important security concern as such de- vices have been modified to allow mobile apps to break out of the normal security sandbox that the OS imposes.		
Missing certificate pinning	Medium	Closed
Certificate pinning is the process of associating the backend server with a particular X509 certificate or public key instead of accepting any certificate signed by a trusted certificate authority. After storing ("pinning") the server certificate or public key, the mobile app will subsequently connect to the known server only. Withdrawing trust from external certificate authorities reduces the attack surface (after all, there are many cases of certificate authorities that have been compromised or tricked into issuing certificates to impostors).		
Stored cross-site scripting (XSS) – Survey answers	High	Closed
The term "Cross-Site Scripting" (XSS) denotes the injection of malicious HTML or JavaScript code into a website. Such an attack is possible if a website receives data from non-trustworthy sources (e.g. URL or POST data) and reflects them directly or indirectly (after storing them to database) to a web browser without prior validation or proper encoding. Thus, XSS allows an attacker to inject and execute code in the browser of his victim. An attacker can use injected HTML or JavaScript code to steal session data and hijack a user session, he can change the layout or behavior of the application or redirect the user to a malicious page. Furthermore, Cross-Site Scripting attacks can be used to exploit browser vulnerabilities.		
Sensitive information in auto-generated screenshots	Low	Low
Android and iOS introduced the screenshot-saving feature for use when the application is sent to the background. This feature may pose a security risk. Sensitive data may be exposed if the user deliberately screenshots the application while sensitive data is displayed.		
Information disclosure – HTTP Responses	Low	Closed
Information like software version numbers, software in use, names of developers or staff, stack traces, file or directory listings, etc. may help adversaries to collect and assemble enough information to enable them to push their attacks further against a system, application, an engineering or staff team. Such information may be contained inside source code comments, HTTP headers, error messages, on status sites, log files, downloadable office document formats (PDF, DOCX, etc.)		

and so on. Every bit of information will help adversaries to prepare more sophisticated attacks against their targets.		
App transport security (ATS) deactivated	Low	Low
App Transport Security (ATS) is a networking security feature by Apple for iOS Apps. It is activated by default and makes sure the app uses secure channels for communications only. The goal is to avoid the usage of protocols and ciphers with known weaknesses by the app. This feature protects the confidentiality and integrity of data in transit.		
Logging of sensitive data	None	Informa-tional
There are many legitimate reasons for creating log files on a mobile device, including keeping track of crashes or errors that are stored locally while the device is offline (so that they can be sent to the ap- p's developer once online), and storing usage statistics. However, logging sensitive data, such as credit card numbers and session information, may expose the data to attackers or malicious applications. Logcat is a system logger, which is always active in the Android OS. Every active mobile application or process can write different types of messages into this log. No special authorization is required for writing messages in the system log under Android. Logcat output can be read out by any other application with the (READLOGS) permission. An attacker can use the behavior to access sensitive data from the application.		

Only one issue remains marked as Medium and unsolved, being the Missing root/Jailbreak detection. However, this is an issue that can impact respondent participation more then it impacts security. Therefore, this issue was set as low priority by Destatis.

17.5 Source Code Analysis MOTUS

The static source code analysis is a static software test procedure, which is assigned to the white-box-test-procedures. In the following, the term "source code analysis" is used as a synonym for Security Source Code Analysis (SSCA), representing the static test of source code for known security vulnerabilities.

The test procedure has been performed by T-Systems Multimedia Solutions (Dresden) using the tool Fortify Static Code Analyzer from MicroFocus™. This tool with automated analyses guarantees an appropriate level of test coverage even for highly complex test objects. The test procedure is divided into following phases:

- Preparation and Information Gathering
- Analyzing
- Information Static Analysis Process
- Final Analysis / Rework / Clean-up

This procedure performs an automated testing of the source code to identify security vulnerabilities.

17.5.1　Test results

[At the moment of testing] A total of 2.724 executable files and 245.131 lines of source code (including third-party libraries) were analyzed with Fortify SCA. Fortify SCA detected 1992 security findings using the most recent rule set. Afterwards, the findings were evaluated and verified by a source code analyst from T-Systems Multimedia Solutions GmbH (False

Positive Analysis). After the False Positive Analysis, the number of identified findings was reduced to 121 actual security-relevant findings.

Table 16: CVSSv3-Rating: test results

Severity	Critical	High	Medium	Low	Informational	Total
Raw Result SSCA	655	141	18	1178	0	1992
After False-Positive Evaluation	32	41	0	48	0	121

The table below gives an overview of the issues and the rating.

Table 17: Source code analysis: issues and rating

Description	Rating	Rating after consultation
Cross-site scripting : DOM (5 issues)	Critical	No issue
The term Cross-Site Scripting (XSS) denotes the injection of malicious HTML or JavaScript code into a website. Such an attack is possible if a website receives data from non-trustworthy sources (e.g. URL or POST data) and reflects them directly or indirectly (after storing them to database) to a web browser without prior validation or proper encoding. Thus, XSS allows an attacker to inject and execute code in the browser of his victim.		
Dangerous function : unsafe regular expression (2 issues)	Critical	No issue (vendor)
Certain functions behave in dangerous ways regardless of how they are used. Functions in this category were often implemented without taking security concerns into account. The split() function is unsafe because it is part of the deprecated family of POSIX regular expression functions which stop reading an input string upon encountering a null byte. Since PHP allows null bytes within strings, the use of these functions can be dangerous in some situations and should be avoided altogether.		
Open Redirect (8 issues)	Critical	No issue
Unvalidated redirects and forwards are possible when a web application accepts untrusted input that could cause the web application to redirect the request to a URL contained within the untrusted input.		
Password management : hardcoded password (1 issue)	Critical	No issue (not used)
Hard-coding sensitive information, such as passwords, server IP addresses, and encryption keys can expose the information to attackers. Anyone who has access to the class files can decompile them and discover the sensitive information which typically leads to a significant authentication failure that can be difficult for the system administrator to detect. Hard-coding sensitive information also increases the need to manage and accommodate changes to the code. For example, changing a hard-coded password in a deployed program may require distribution of a patch.		
Path manipulation (2 issues)	Critical	No issue
A directory traversal attack (also known as path traversal) is vulnerability that allows an attacker to read arbitrary files on the server that is running an application. The vulnerability occurs if the application uses external input to construct a pathname that is intended to identify a file or directory that		

is located underneath a restricted parent directory, but the software does not properly neutralize special elements within the pathname that can cause the pathname to resolve to a location that is outside of the restricted directory. By manipulating variables that reference files with "dot-dot-slash (../)" sequences and its variations or by using absolute file paths, it may be possible to access arbitrary files and directories stored on file system including application source code or configuration and critical system files.		
Privacy violation (10 issues)	Critical	No issue (vendor)
Privacy violations occur when: 1. Private user information enters the program. 2. The data is written to an external location, such as the console, file system, or network.		
Privacy violation : shoulder surfing (4 issues)	Critical	No issue or fixed
A password need not be viewable to its owner, and must not be viewable to others. If a password is displayed in plain text, anyone in the vicinity could see and use it to compromise the system. In computer security, shoulder surfing refers to using direct observation techniques, such as looking over someone's shoulder, to get information. Shoulder surfing is particularly effective in crowded, public environments. This threat particularly applies to mobile devices, which are generally intended for use in all environments, both private and public.		
Cookie security : persistent session cookie (1 issue)	High	No issue (not used)
A persistent session cookie remains valid even after a user closes his browser and is often used as part of a "Remember Me" feature. Consequently, a persistent session cookie allows users to remain authenticated to an application even after closing their browsers - assuming they didn't explicitly log out. This means the next person that opens the browser will automatically be logged in as the last user. Unless your application is deployed in a controlled environment where users are not allowed to log on from shared machines, it is possible for attackers to compromise your user's accounts even after they've closed their browsers.		
Denial of service : regular expression (1 issue)	High	No issue (fixed/not used)
A Denial of Service (DoS) attack tries to overload a system or network with the goal to limit the availability to its regular users and thereby inflicting damage because the system is not available. Targeted resources are a service's CPU, volatile memory (RAM), persistent memory (hard disk), network bandwidth or socket count. A DoS attack may inflict temporal disruption (by the end of the attack, the service returns to its normal state) or persistent disruption (services are stuck in an overload state, even after the attack has ended; accounts may be locked, storage be depleted). DoS attacks might be used to camouflage other attacks.		

Insecure randomness (35 issues)		High	Fixed
	When software generates predictable values in a context requiring unpredictability, it may be possible for an attacker to guess the next value that will be generated, and use this guess to impersonate another user or access sensitive information.		
Object injection (4 issues)		High	No issue (not used)
	Object injection vulnerabilities occur when untrusted data is not properly sanitized before being passed to the unserialize() function. Attackers could pass specially crafted serialized strings to a vulnerable unserialize() call, resulting in an arbitrary PHP object(s) injection into the application scope. The severity of this vulnerability depends on the classes available in the application scope. Classes implementing PHP magic method such as __wakeup or __destruct will be interesting for the attackers since they will be able to execute the code within these methods.		
Command injection (2 issues)		Low	No issue (not used)
	Command injection vulnerabilities take two forms: • An attacker can change the command that the program executes: the attacker explicitly controls what the command is. • An attacker can change the environment in which the command executes: the attacker implicitly controls what the command means. In this case, we are primarily concerned with the second scenario, in which an attacker can change the meaning of the command by changing an environment variable or by inserting a malicious executable early on the search path. Command injection vulnerabilities of this type occur when: • An attacker modifies an application's environment. • The application executes a command without specifying an absolute path or verifying the binary being executed. • By executing the command, the application gives an attacker a privilege or capability that the attacker would not otherwise have.		
Cross-site request forgery (13 issues)		Low	No issue or fixed
	Cross-Site Request Forgery (CSRF) is an attack where a valid request to the server is being invoked by a remote attacker in the victim's browser session. A precondition for this attack is that the victim must be logged into the web application. If a victim can be forced to visit an attacker-controlled page which contains, for example, a JavaScript code snippet with a POST request on the vulnerable page, the action will be executed with the session that the user has on the vulnerable web application.		
Password management : password in comment (1 issue)		Low	No issue (not used)
	It is never a good idea to hardcode a password. Storing password details within comments is equivalent to hardcoding passwords. Not only does it allow all of the project's developers to view the password, it also makes fixing the problem extremely difficult. After the code is in		

production, the password is now leaked to the outside world and cannot be protected or changed without patching the software. If the account protected by the password is compromised, the owners of the system must choose between security and availability.			
System information leak : external (17 issues)		Low	No issue
Information like software version numbers, software in use, names of developers or staff, stack traces, file or directory listings, etc. may help adversaries to collect and assemble enough information to enable them to push their attacks further against a system, application, an engineering or staff team. Such information may be contained inside source code comments, HTTP headers, error messages, on status sites, log files, downloadable office document formats (PDF, DOCX, etc.) and so on. Every bit of information will help adversaries to prepare more sophisticated attacks against their targets.			
Weak cryptography hash (15 issues)		Low	No issue or fixed
MD2, MD4, MD5, RIPEMD-160, and SHA-1 are popular cryptographic hash algorithms often used to verify the integrity of messages and other data. However, as recent cryptanalysis research has revealed fundamental weaknesses in these algorithms, they should no longer be used within security-critical contexts. ⬚Effective techniques for breaking MD and RIPEMD hashes are widely available, so those algorithms should not be relied upon for security. In the case of SHA-1, current techniques still require a significant amount of computational power and are more difficult to implement. However, attackers have found the Achilles' heel for the algorithm, and techniques for breaking it will likely lead to the discovery of even faster attacks.			

After a one-week period of mitigation a consultation between hbits and T-System took place evaluating whether the issue was a 'no issue' or was fixed. The last column gives an overview of the status closing all issues. This status was accepted by Destatis.

18 WP5 – D5.4 - Update CSPA-catalogue

Overall goal: update of the CSPA Inventory of Eurostat.

The CSPA-catalogue related to MOTUS and HBS can be found here:
https://webgate.ec.europa.eu/fpfis/wikis/display/ISTLCS/HBS+TOOLS+MENU

19 References

Bevans, G. (1913). How working men spend their spare time. *Columbia University Press*.

Chinosi, M., & Trombetta, A. (2012). BPMN: An introduction to the standard. *Computer Standards & Interfaces*, 124-134.

Commission, E. (2008). Guidelines on Harmonized European Time Use Surveys. *Eurostat*.

DGINS. (2011). *Wiesbaden Memorandum*. Opgehaald van https://circabc.europa.eu/: https://circabc.europa.eu/sd/a/51a4bcbd-2ac2-46a8-8992-cfb9e6009522/Item_3.3.%20Modernisation%20of%20Social%20statistics_annex.pdf

Eurostat. (2020). *Harmonised European Time Use Surveys (HETUS) — 2018 guidelines — Re-edition* . Opgehaald van https://ec.europa.eu/eurostat/web/products-manuals-and-guidelines: https://ec.europa.eu/eurostat/web/products-manuals-and-guidelines/-/ks-gq-20-011

Eurostat. (2021, 03 11). *Inventory of tools and sources for diary-based data collections*. Opgehaald van ISTLCS/INVENTORY: https://webgate.ec.europa.eu/fpfis/wikis/display/ISTLCS/INVENTORY

Eurostat. (2022, 09 06). *Household Budget Surveys (HBS) - Overview*. Opgehaald van Eurostat, Your key to European statistics: https://ec.europa.eu/eurostat/web/household-budget-surveys

Harvey, A. (1993). Guidelines for time use data collection. *Social Indicators Research*, 197-228.

Hoorn, E., & Montagner, C. (2018, November). *https://www.rug.nl/research/research-data-management*. Opgehaald van Starting with a DPIA methodolgy for human subject research v. 0.1: https://www.rug.nl/research/research-data-management/downloads/c2-dataprotection-dl/dpia_guidance_doc_v1_pub.pdf

Juster, T., & Stafford, F. (1991). The allocation of time - Empirical findings, behavioural models and problems of measurement. *Journal of Economic Literature*, 471-522.

Minnen, J. E. (2020). *SOURCE TM: Software Outreach and Redefinition to Collect E-data Through MOTUS. Towards a Modular Online Time Use Survey*. Brussels: Statistics Belgium, Destatis, hbits CV and Vrije Universiteit Brussel. - TOR.

Nations, U. (2004). Guide to producing statistics on time use. *Statistic Devision, Department of Economic and Social Affairs*.

Pember, R. (1914). Round about a pound a week. *Garland Publishing*.

Stinson, L. (1999). Measuring how people spend their time - A time use survey design. *Monthly Labour Review*, 12-19.

Szalai, A. (1972). The use of time - daily activities of urban and suburban populations in twelve countries. *Mouton*.

The European parliament and the council of the European Union. (2019, 10 14). *Regulation (EU) 2019/1700 of the European Parliament and of the council of 10 October 2019 establishing a common framework for European statistics relating to persons and households*. Opgehaald van https://eur-lex.europa.eu/: https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=uriserv:OJ.L_.2019.261.01.0001.01.ENG

Zuzanek, J. (1980). Work and leisure in the Soviet Union - A time budget analysis. *Praeger*.

Annexes

Annex 1: WP4 – D4.1 – UI/UX for TUS and HBS

Presentation to accompany the UI/UX slides at the beginning of the project:

- B5460_BE_4.1_UI-UX_20210702_Presentation_UI_UX_MOTUS

UI/UX slides MOTUS:

- B5460_BE_4.1_UI-UX_MOTUS_ONBOARDING_overview.pdf
- B5460_BE_4.1_UI-UX_MOTUS_MENU_overview.pdf
- B5460_BE_4.1_UI-UX_MOTUS_MY_GROUP_overview.pdf

UI/UX slides HBS:

- B5460_BE_4.1_UI-UX_HBS_All_ExpenseLocations.pdf
- B5460_BE_4.1_UI-UX_HBS_homepage.pdf
- B5460_BE_4.1_UI-UX_HBS_new_expense_add.pdf
- B5460_BE_4.1_UI-UX_HBS_new_expense_place.pdf
- B5460_BE_4.1_UI-UX_HBS_new_expense_ticket_scanning.pdf
- B5460_BE_4.1_UI-UX_HBS_homepage_edit_expense.pdf
- B5460_BE_4.1_UI-UX_HBS_Shoplist_GER.pdf
- B5460_BE_4.1_UI-UX_HBS_COICOP_9_categories.pdf
- B5460_BE_4.1_UI-UX_HBS_ECOICOP_FIXED_COSTS.pdf

UI/UX slides TUS:

- B5460_BE_4.1_UI-UX_TUS_TIMELINE_overview.pdf
- B5460_BE_4.1_UI-UX_TUS_NEW_ENTRY.pdf
- B5460_BE_4.1_UI-UX_TUS_FUNCTION_DESIGN.pdf
- B5460_BE_4.1_UI-UX_TUS_TL_SHARING_overview..pdf
- B5460_BE_4.1_UI-UX_TUS_TENTATIVE_LOCATIONS_DESIGN.pdf
- B5460_BE_4.1_UI-UX_TUS_SEPARATE_DAYS.pdf

HBS concept overview

The table on the next page gives an overview of the different overviews for HBS and was constructed in May 2021. Each title corresponds to the name in the folder. The general comments are on the last page.

Sometimes the word 'shop' or 'product' is used. This is a very silly generalization but serves to make it more understandable. (For example, if there is a supermarket in the designs).

HBS Homepage

1	Landing page standard
	• Calendar is sticky, stays on the screen when scrolling.
	• (Scroll down) list of expenses limited by day.
	• 1A: default view of landing page. With calendar unfolded.
	• 1B: calendar folded
	• 1C: Detail of an edition with options. Edit, delete and more (**See HBS HOMEPAGE -Edit/copy/share/extra functions.pfd**)
2	Tentative dates
	• Similar to TUS in shape and placement.
3	Commited + Tentative
	• 3A similar must TUS in form and placement.
	• 3B. Also options for add and source similar to TUS
4	Indication of incomplete data
	• Red line in Calendar (similar to TUS)
	• The relevant edition has the text 'Please add additional info'. The less red the better. But if necessary, this one will be red.
5	'New expense' start screen
	• 4 options that the respondent first receives when a new issue is published.
	• Scan ticket: complete processing of the entire purchase ticket
	• New expense: manual entry of expenses (See the pdf series **HBS New expense**)
	• Return: Indicate immediately that you have returned something to the shop.
	• No expenses today. Indicate that you had no expenses that day. If you have already entered expenses for this day, this button is greyed out.
6	No expenses
	• The respondent has indicated for this day that he did not incur any expenses. **(Screen 5)**
7	Menu integration
(8)	Additional possibilities
	• The landing page can contain any content. A summary of the statistics (with button to full statistics), help functions or tips, etc.

STATBEL
Belgium in figures

D STATIS
wissen.nutzen.

VUB VRIJE UNIVERSITEIT BRUSSEL

h

eurostat

HBS NEW EXPENSE: Product input/search/filtering

This PDF only deals with the button 'Add a new expense or return shipment' when entering a new expense. This is a broad representation of the concept. How to question different types of expenditure may require a more individual approach.

0	Start. New expense button off **HBS homepage.pdf**
1	Main screen New expense • Fill in what can already be filled in (country, date) • If date should include time, decide whether to fill it in or not. (Otherwise we will immediately get an error message saying that it is only 'half filled in'), unless we give the time at which the respondent enters the purchase.) • What was your expense: the full list of purchases in this shop comes here.
2	Start screen new expense • Start of new expense. No categories, previous expenses, favourites, suggestions, etc. Respondent can type or scan the barcode. • Filters are all on by default with every new release? • Scan product/barcode button.
3	Respondent starts typing • COICOP results start filtering. (Currently up to A) • Respondent does not type out his expense completely, but starts filtering the results.
4	Filters • 4A. All filters are always on by default. Respondent can switch them off or on again. • 4B. Filters off = grey-blue. The orange dot (light) is off. • With 'Set' or < button respondent goes back.
5	Filtered results. Respondent can now select its COICOP and confirm with 'Done'.
6	Additional info on the publication • 6A. The extra info according to the release is already in place as best as possible. (i.e. for t-shirts the app knows that the unit will probably be 'pieces' and not 'kg'. • 6B. The respondent can complete the name of the expense within this screen. If he presses COICOP, he will go back to **Screen 5** • 6C + D. Units (piece, %,...) can be changed with pop-up screen. Still thinking how to properly shorten 'free item'. • **Note: How far do we go in discounts? %, freebie, 1+1, 2+1,... Albert Heijn already does 1+6, or 'x+x' and then fill in the rest yourself?** • 6E. Respondent has filled in everything correctly and can now press 'add'.
7 (=1)	Product has been added, together with some other products. Respondent can: • Add extra products via 'Add new expense or return shipment'. • Make adjustments to the product by pressing. It then returns to **Screen 6E.** • Answer the overage questions and press Save, after which the edition is added to his home screen (Calendar)
BIS	Barcode scanning • Instead of adding via COICOP, the respondent can also choose to scan the product's barcode in **Screen 2**. • BIS 1. Scanner opens. If the barcode is read correctly, the screen automatically switches to **BIS 3.** • BIS 2. Respondent can also choose to enter the barcode manually. If he presses 'done', the screen automatically goes to **BIS 3**. • BIS 3. The result of the scan. The app recognises the barcode and can attach the correct COICOP (as well as unit). The respondent still needs to give the product a name here (unless the app knows this name too) From **BIS 3** onwards the respondent follows the same route as **Screen 6A**

Extra question: Can the respondent also just type in the COICOP code in screen 2 and choose his result? For example, he is going to buy a white bread (01113A) every day at Wendy's Bakery and just enters 01113A.

Answer: Of course, the result will then immediately appear filtered. In screen 6, the name edition will be 01113A (which is what he typed), so he can change the name there to 'bread'.

But: What will be much easier for the respondent in this case, is just to copy the complete expense in the home screen to the next day. If he has bought extra coffee cakes that day, he can easily add them when copying.

HBS NEW EXPENSE: Place of issue

This overview only deals with entering the place of publication.

0	Start. New expense button off 1. HBS homepage.pdf
1	Main screen New expense • Fill in what can already be filled in (country, date) • If date should also include time, decide whether to fill it in or not. (Otherwise we will immediately get an error message saying that it is only 'half filled in'), unless we give the time when the respondent enters the purchase as the hour). • What was your expense: the full list of purchases in this shop comes here.
2	Start screen 'Where did you make the expense • Respondent enters name of shop/service provider/... • Button suggestions: based on tentative data. • Previous: for quick access to a place that the respondent frequents (a certain supermarket, website, ...).
3	Respondent starts typing • List of possible locations = Name of the shop and therefore not 'Carrefour Wilrijk, Carrefour Boom, Carrefour Aartselaar,...) • Add new' at the bottom of the list. Option to add place if no result matches what respondent means. • Done' => respondent goes on to **screen 5.** • Add new' respondent proceeds to **screen 4B** (where 'Carrefour' remains as name.
4	Shop not found/add • 4A. The shop is not in the database. Respondent can add shop to database. (Respondent will first check if he did not make a typo). • 4B. Name remains. Respondent is asked to select a main category. When selected, automatically goes to Screen 4C, Add button for category selection has green warning. • 4C Respondent is asked to select subcategory. Add button for category selection has green warning. • 4D Respondent activates subcategory. Enable multiple options? (Local shop) He can now press Add and goes to **screen 5.** • (4E) all icons in colour for reference.
5	Shop is filled in in the main screen of New expense. Respondent can fill in further and press save to add the expense to his diary.

HBS HOMEPAGE: additional functions

This table gives an overview of Edit, Delete, Search, Copy, Split and Share.

1	Home page.
	• 1A opens the options: Delete, Edit, and the knock with extra functions (Analogy TUS).
	• 1B: Additional functions button.
	• Copy: copy the entire edition (to a new day). Useful if you go to the same bakery every day to buy fresh bread.
	• Split. Split your expenses. Split also allows you to easily add the same shop on the same day with an empty shopping basket. Handy if, for example, you forgot one thing in the supermarket and went back quickly. See **screen 5**
	• Share: Share with 'My Party'.
	• Note: unlike TUS, **there is no insert**. Since an issue is determined at the time of release, you cannot split that infinite millisecond and quickly do another issue in between.
2	Delete
	• Analogy with TUS. Pop-up to confirm action.
	• Only if there is no expenditure that day will a message appear (**screen 2B**). (So there are no "Gaps" like in TUS).
3	Edit a purchase
	• 3A. Analogy TUS: press the 'entry', expense and you will see your fully completed expense.
	• Here you can change answers, add new editions or delete them.
	• 3B. By clicking on an issue, you will see the completed questions of that issue and you can also edit them.
4	Copy
	• Copy is useful if you regularly make the same expenses. For example, going to the bakery every day.
	• As with TUS, all you have to do is enter a new date on the copy.
	• Of course, all other questions can also be adapted. Did the respondent also buy coffee cakes that day, or brown bread instead of white? Then these can be adjusted quickly.
	• Note: If you were to copy to the same day, this is also a way of adding, for example, that quick trip to the supermarket for that one forgotten item. However, you then have to delete all other expenses entered one by one. **Split (screen 5)** is a faster solution for this.
5	Split
	• Split can be used for several things:
	• You have shared your purchases and want to indicate this split. (For example, to share a piece with My Party afterwards, or you have a ticket in Mediamarkt for 5 things of which you have had a professional invoice made for 3).
	• Split can serve as a quick way to duplicate the entire entry with an empty shopping basket. Useful if you have run back inside after shopping in a supermarket because you forgot one item and have a separate ticket.
	• is going to be especially useful if you can scan and enter a ticket in a few clicks. 2. Is more convenient for manual entry.
	• 5A: Pop-up screen shows the list of expenses/returns and asks you to select only those you wish to set aside. If you do not select any, the entry is copied with an 'empty basket'.
	• 5B. Bananas and hair dryer selected. The issue has been split.
	• 5C. No items (bananas, hairdryer) selected, so duplicate with 'empty basket'. Respondent is prompted that data is missing.
6	Share

	• Like TUS, you can share entries (expenses) with 'My Party'.
	• 6A. Up to here, it is standard analogy with TUS. You share your entry with a Party member. After this, it can be completed.
	• 6B. Idea of having extra to indicate how the costs are distributed. Every fifty-fifty, etc.
	• 6C. expense is shared. Adjustment of the amount if a sharing has been opted for? In the details you will then see the total amount, the division and with whom shared. (Total amounts will then also be adjusted.
7	Search
	• Analogy with TUS. All expenditures that match the search term are shown on the screen in scroll.

Notes

1. We are going to have to make a clear verbal distinction between an 'expense'(=entry) in the homescreen; and an 'expense'(=product/service/...) in as part of the list of expenses within that expense (New expense).
2. Are there alternatives to 'Split'? If splitting purchases is not necessary, only 'copy the place of issue but not the issue itself' remains. Maybe it can be simpler then.
   • Put a button in the 'New expense' screen to 'delete the entire contents of your basket in one go'?
   • Replace 'Split' with another button that says you only make the full copy, but with an empty 'basket'. (Disadvantage: long name, ugly button).
   • Advantage of Split: We stay within the functionality of MOTUS without adding a new one.
   • Disadvantage 'split', in publishing terms this rather refers to 'splitting the bill with someone'.
3. How far do we go in discounts? %, freebie, 1+1, 2+1,... Albert Heijn already does 1+6. Or a form like 'x+x' or 'x+1' be clear to the respondent and then fill it in yourself?
4. Make it clear to the respondent that 3 free toothbrushes with 1 tube of toothpaste is not '1+3' but 3 'free items' that must be entered separately.

The concept overview for HBS and UI/UX for TUS and HBS were presented during the first general meeting on July 2nd, 2021.

After the presentation of the concept overview for HBS and UI/UX for TUS and HBS, feedback was provided by both Destatis and Statbel. The feedback went back and forth 2 times.

Feedback on the MOTUS interface for TUS & HBS

The feedback is based on the Mockup presented in the meeting 2nd of July 2021.

- The comments in black are the additional comments.
- The comments in green are the responses from hbits.
- The comments in orange are the second round of feedback from Destatis.

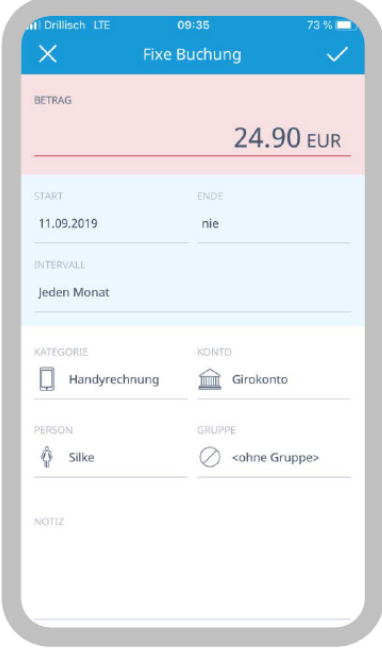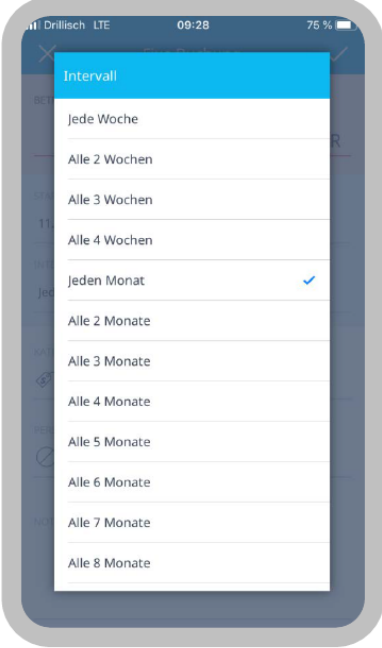| Reference (screen number)	Comment
General – Icon labeling	• We recommend adding small text labels to the icons, at least the ones that are rather unique to this app or generally not commonly used (e.g. icons such as the ones the taskbar at the homescreen, the source (group/gps) icon for tentative data, etc.) – Both UX guidelines & our pretest results attest a better usability when adding text labels. For some icons, it might even be better to just show text, without a visual. • Fully agree. Adding text labels to unique icons is definitely a thing we consider at every turn. New expenditure, copy-split-duplicate, the filters,…. The 4 bottom buttons of the UI however, are basically on and off switches and are designed to be very intuitive. E.g.: tap on the magnifying glass, the search bar (with 'search' pre-filled pops up. Tap on it again, it disappears.) By going with this intuitive way of handling these buttons, we clear of space and reduce text clutter. • We understand that and can keep it like that for now, albeit I'm still a bit sceptical if all users will understand these. • A good example of where it is already implemented like this is when adding a new expenditure – all icons feature text labels. • Ideally, all text labels should clearly communicate the underlying action, e.g. "" instead of just "return" (shown when clicking the "plus"-button for a new expenditure). • This is true. It should always be clear. At the end of all our UI and UX designs, we always bring a copywriter aboard to dot all the i's in this.
General – Statistics	• As statistics about one's own expenditures serve as an important intrinsic motivation to participate in the HBS, they should be accessible directly / easy to navigate to (e.g. as one of the main options on the bottom overlay taskbar) • Currently, one has to scroll down on the homepage to see it, correct? (1. HBS homepage, #8) – some will likely not notice it like this. Also, it requires more user interaction to get to it, especially if one has entered many expenditures for a given day – in that case, one has to scroll down far to reach the statistics • This is something to think about. Either we put a button there. I would suggest next to daily total and grand total. Or we leave it like it is. The reason the home screen is designed the way it is, is because to me, it consists of widgets. The calendar + daily expenses and grand total is a widget, the expenses per day is one too. The stats are one as well. • By making them widgets, we can already give some main info on the home screen without taking up all the space. And the full info can he either expanded or visited on a different page if the respondent is thirsty for it. Not only would the information be either everything or nothing if we were to make it a button, as a

	widget it also explains itself by showing the content, whereas a button would need explanation. And lastly, it's a trend that you clearly see in many app updates since 2020, and is still ongoing in 2021.
	• Ok, we can follow and agree.
	• Remarks on this: On the expense list widget. We could either make this an expandable widget, only showing the most recently logged expenses of that day, with a 'show more' button. But we could also consider that on an average day, this might not be such a long scroll.
	• First option seems to me an appropriate solution, but I can't make a picture of what you mean with the second option and "an average day". Could you explain bit more?
	• If not already implemented: Should ideally feature averages and within-subject comparisons (e.g. week 1 vs. 2, month 1 vs. 2, individual timeframes, etc.) and – in a first best solution – the possibility to export data as an excel-sheet
	• The statistics page is still a concept only, and will be one of the last items on the design list. I feel that first, we need to design the app in a way the respondent can easily log his myriad of expenses, and after that we can focus on how to display the statistics around them. But it's definitely on the list and of course all suggestions in this are welcome.
	• Is this only true for the CRŒSS project or could it come true (i.e. a "must") for the EVS project then (if solution for data security – next point)?
	• Be aware that, when it comes to displaying statistics, we'll also need to take into account the rules that data can't be stored on the advice, and other possible restrictions on the data security level.
	• Correct.
General – Menu	• In HBS the time-factor is not as relevant as it is in TUS. The flow might be limiting in terms of what users can edit in later stages of their participation. Especially since a longer period needs to be covered, users should be allowed to make changes to their "diary" (spending) in retrospect and in their questionnaires during the completely collecting period.
	Hence, maybe we again should discuss the opportunity to place the sub-sections (states, questionnaires) in the burger menu and make them there reachable.
	From my designer point-of-view: any changes to their diary are as simple as in TUS. Simply navigate to the day/expense or search for it with the search function, tap it, and press edit. So this is definitely fully implemented.
	The different tasks should not be placed in the burger menu as separate sections. It can be a long list, removing the power of the menu. All tasks can be found under the My research tab.
	Right, I guess it's not that much the design question, but the data editing (see below)
	As for editing the questionnaires we again face the sequential flow of actions. It remains a difficult element.
	Difficult, but managable, right?
General – Taskbar (homescreen)	• It is not intuitive at first what purpose the four main features on the bottom taskbar each serve
	• The 4 buttons are intuitive on/off switches. If you press them, immediately something happens without the user leaving the

	screen. The button itself lights up green (on-mode). To turn it off, simply tap again.
	• Realisation in app will show intuitivity and has to be proofed by pretest (CRŒSS/mainly in EVS, as with external respondents).
	• The bottom buttons are important elements for both TUS and HBS. Visual consistency here is key to us.
	o The Tutorial icon on the bottom left – as it gives information instead of having the user interact/actively add data, it might be a good idea to separate it visually & move it away from the interaction-icons, for example to the top right of each page (and remove the shopping cart-icon which has no underlying functionality). This way, it can also be implemented on all sub-pages, where currently no taskbar on the bottom exists. Respondents will learn this and then intuitively know where to look for information, no matter on what subpage they currently are. The Assistant is more than a tutorial, it's an assistant. (Incidentally, it can also be toggles on an off from the help page in the menu. Though this button is of course quicker and easier to navigate.) For visual balance as well as consistency, the task bar will remain in place the way it is and will be evaluated at a point in the future. Assistant to be tested in the CRŒSS pretests.
	o The search-icon: Unclear, what exactly is being "searched" with the icon on the left. (Search through (all?) my expenditures?) It's a good point that will be easily fixed with the default text in the search bar. When you tap the icon, the search bar opens and there is a thin, italic default text (Something like 'Start typing what you want to search for.') The copywriter will later put a far better line in here, I'm sure. My aspiration is you can search for anything. So you can type a store name, but also a category and even an individual expense. So let's say if I were to type 'cheese', all the store/expense location results in would be places where I had bought cheese in. So the default text could say something like 'Type any location, category or expense' (Or whatever better solution the copywriter provides)
	o For design ok. Also needs to be evaluated how good the search-function really is …
	o Is the "group" feature really one of the main options, is it used that often? If not, might place it somewhere less prominent (e.g. menu) The group feature is a key ingredient in researches and will become even more important in the future. Especially in household budget surveys. However, this button has always been implemented as a 'need only' button. Which mean that in researches where the Household is not a factor (e.g.: certain TUS researches), this button will also not show up. Not even grayed out, it simply won't be there so it won't cause confusion.
	o Agreed.
	• Icon on the far right: Unintuitive at first what the icon does. Also again to be considered, is the functionality used often enough for it to be placed that prominently, or could it be implemented somewhere else? See above. On/off switch in this case with 3 modes. As for tentative data, this is also one of the key features of the MOTUS framework.

HBS homepage – #1: incomplete/ entered data	• Maybe work with small icons within each calendar day to show whether a respondent already enters data/expenditures (or confirms that he/she had no expenses). This is a really good idea. Will look into an implementation for this. At the very least, color difference in filled out or not filled out, showing the respondent's progress, is desirable. • Agreed.
HBS homepage – #1A: aggregate periods	• For Germany the collecting period will be 3 months, therefore maybe total sum (Grand total) should be also be possible to show for month (e.g. June total). Maybe some kind of a "zoom out" function is a solution for it, aggregating the calendar from days to weeks to months. We can look into this. In MOTUS, all white rectangle are pressable buttons/interactions. So the info of daily total/grand total has been designed with the idea to leave the option to add something around it. (Incidentally, should be decide on not putting anything behind these buttons, the design of their numbers will change to something visually non-interactive) When it comes to monthly total, of course there is the question how often this info will be correct. • Would it be possible to even split these 2 to have 3 sums (i.e. single day, actual month, Total spendings?). Or single day and actual month next to each other, and below the grand total? Only ideas ... • See next comment. • If a user started his month on the 15th, does the monthly total give accurate information? • In Germany we will have 4 quarters of the year, thus they all start the beginning of the month. • Another indication is when a user stops his research and is being replaced by a new one on the 27th of the month, again it would only give little info • 'Grand total' simply indicated: this is how much you have spent since you joined the research. • Stats for Germany should show spendings for each months at any rate. • And or course daily total: a user can start mid-month, however he can't start mid-day.
HBS homepage – #1B/C & 3A/B: editing entries	• Is it intuitive for respondents to click on expenses to see more details/change data? Maybe an item (e.g. arrow) can give a hint to the users that editing is possible. Everything on a white BG is a button. At the start of the design of the MOTUS framework, we choose not to add collapsible arrows to keep the design less cluttered. To our opinion, apps have been part of our lives long enough that users intuitively press on buttons. Within the MOTUS framework, everything on-screen that is in a white rectangle, should be pressable/interactive of have an interactive icon on it. If it's text on a grey background (such as Expenses, Wednesday June 17) it's a title and not pressable/interactive. • Agreed. Guess, the assistant should make it clear. Anyway, we will see by testing.
HBS new expense – general: fixed expenditures	• Maybe that is a special German requirement as we have the long reporting period of three months, but in our opinion solutions for entering fixed/repeated expenditures (e.g. housing rents) is missing. One opportunity could be to add a button/field at the screen for new expenditures (HBS new expense-uitgave invoeren-zoeken-filteren #1) pointing out the frequency the expenditures

appear (see "Zeitintervall" at the exemplary screenshot from another mobile app). Another one could already separating it when clicking on the plus-button by adding "Fixed expenses" to the selectable alternatives.

- Technically these expenditures should be pointed out in the calendar on every day given the information added by the respondents (e.g. every 1st of the three reported months) and accordingly stored in the database (three times). I completely agree. A fixed cost needs to be added. Behind the + button is indeed the obvious choice. Since this type of expense required different details, it should serve as a design of its own.
  Agreed.

Fixe Buchung	**Intervall**
BETRAG: 24.90 EUR	Jede Woche
START: 11.09.2019 ENDE: nie	Alle 2 Wochen
INTERVALL: Jeden Monat	Alle 3 Wochen
KATEGORIE: Handyrechnung KONTO: Girokonto	Alle 4 Wochen
PERSON: Silke GRUPPE: <ohne Gruppe>	Jeden Monat ✓
NOTIZ	Alle 2 Monate, Alle 3 Monate, Alle 4 Monate, Alle 5 Monate, Alle 6 Monate, Alle 7 Monate, Alle 8 Monate

---

HBS new expense-uitgave invoeren-zoeken-filteren – #4A/B

- "Filter" option:
  - Functionality: while having pictures on a new page looks great, a probably easier way to use them would be to have only text labels and have them scroll out on the same page when clicking the filter button. That way, one stays on the same page and can dis/enable product type by product type more easily, depending on the search results found. This could save several clicks and make it easier for the user. Yes and no. Putting everything on it, makes it a list with categories. Like expense, or 'activity on TUS') Yes, lists would be easier (definitely if the list of categories would be too long), but since adding an expense in HBS is a longer procedure than e.g. adding an activity in TUS, choosing for icons spruces up the design and user experience. It also serves as a landmark on the roadmap that is the app, (one cannot visually mix it up with expense lists) and icons add info which makes these sometimes rather vague categories easier to understand. Since we can add 3 or more specific illustrations in the icon we can make something like 'personal services' more understandable.
  - Ok, agreed for now. Implementation and tests will show, if the double click will work. Experiences have shown, that two-clicks (e.g. for branches of trade) were rated negatively,

	as respondents always had to go back again. However, we see also your argument for icons as complexity is higher.
	o Design: As it supports the main search, rather than being an alternative to it (such as the scan product barcode), it might be more fitting to display it differently, to show that it acts as a supplementary to the search bar on the top I agree. Will look into it to make it more standalone.
	o Name: Could name it "Filter Categories" to directly show what type of filter is meant, as there is (as far as we can see) no other, second filter available. That could already do the trick, indeed.

HBS new expense-uitgave invoeren-zoeken-filteren – #6D/E	• Discount:
	  o Just in case this is not hidden behind "see more" – it should be possible to enter discount in absolute € (e.g. "0,50€"), not only in % (or for free). Agree, and we should put it before see more as well. Definitely that pop-up had room for more than 2 options before hitting see more.
	  o Could be a bit clearer that the discount is already calculated in the overall price. (Maybe simply add "Discount included" instead of only "discount".) Or 'product discount' versus 'overall discount'? Good point and something to flag to our copywriter.

HBS new expense – ticket scanning – #2/3	• Is there an app-sided check planned, that evaluates whether the quality of the photo is sufficient and if not – gives immediate feedback to the user? That would make it possible to have the user take another photo to try and correct possibly blurry or unclear images. Yes there should be. Same as in panorama, the app should guide the user. Should come back from microservice.
HBS new expense – ticket scanning – #3	• Idea: The "save"-button could be greyed out until one has scrolled down and seen all automatically captured expenditures – to make sure one checks all entries for mistakes / wrong expenditures, similar to how accepting the terms of use work on some apps. Good idea, definitely a will-do.
HBS homepage - uitgave bewerken – #1B	• The popup with textual actions that appears after clicking on the three dots is great and easy to understand. However, the three dots themselves on the left were overseen by several test persons in our current German MOTUS Test – an idea to simplify it is to implement the popup when clicking on the "edit pen" on the right instead. Clicking the pen would open a popup, with "Edit" on top, followed by the three other options. That would make sure that the additional options are seen and it would allow to remove the three dot icon altogether, simplifying the design. That could work, but would need implementation on TUS as well. It would depend on what people use the added functions most for. If 90% of the time people will just want to edit or delete, then making these buttons quick helps the user experience. But if it were to be more '50-50', I agree it has value to put it behind 1 pencil button. However, the fact that they were overseen, is definitely something to take into account. Thanks for the feedback on this. • As we intend for ZVE to fix only findings with high impact (due to cost/timing reasons), this is something for the longer list, but could be already taken into account for HBS.
HBS homepage - uitgave	• Currently, in order to add "no expenses", one has to click on the plus to actually "add expenses" first, only then can one confirm no expenses. This is somewhat counterintuitive. Depends on how you see it. Is the plus button a button to add an expense, or a button to

bewerken – #2B	add a data entry to the research? To us, it's data. By moving everything behind 1 plus button, we centralize the UX, and remove clutter. • Sounds reasonable. • An idea to change the way of how adding "no expenses" works: The second part of the info text saying "Please add an expense or confirm that you made no expenses on this day." could feature a direct link to a confirmation popup, visualized by e.g. underlining "… confirm that you made no expenses". That would save users a click and might make it more intuitive. • The user can also now confirm that he made no expenses this day with 2 button presses. + and "no expenses". Should we put this behind the 'new expanse button', we need 3 presses. A user who made no expenses on his very first day, would also only learn about this option the first time he made an expense. And would be stuck with a notification of incomplete data on his first day. • In conclusion, we advise on keeping "no expenses' where they are. • Agreed

Annex 2: WP4 – D4.2 – Report on the usability tests from Destatis

Report by Destatis – Sect. B 22 - Daniel Knapp, Christiane Richter, Anja Sommer, Pablo Brecht

General Usability and Layout

TP liked the general layout and appearance of the app. Nonetheless, some issues should be resolved that currently reduce the user-friendliness (e.g., one TP mentioned it "does not feel ready yet").

Interaction and navigation

- Button display:
  - The order of primary and secondary functions should be kept constant across the app: "back"/"quit" on the left side and "ok"/"continue" on the right.
  - Further, the "cancel"-button is sometimes used rather as a "back" button, as it leads back to the last process of entering and not the whole set is deleted.
  - On several pages, the button to "add" an item is covered by either the keyboard or other elements.
- Scrolling was possible on a few pages of the app. In some cases, this was not indicative from the design per se, e.g. when content on the bottom of the page was cut off in a way that it looked like there was nothing else on the page. A visual cue is missing whenever scrolling is possible.
- When entering numbers, instead of the keyboard that includes an array of special characters and numbers, a sole number pad should open, as no other characters are needed here.
- As decimal separators are commas in Germany and dots in English speaking countries, they should both be accepted or be automatically converted.

Design

- The two colored dots (that only serve a design purpose but have no inherent function) located at the top and bottom of nearly all pages of the app confused the test users. These imply the existence of two pages, navigable via swiping, as this is what users have learned from other apps (and even from the login pages of this app). However, this is not given:


- Some currently small icons could be slightly larger to increase tap efficiency.

Login & Welcome pages

All in all, the test showed that the login and welcome pages contain some bugs and input difficulties that limit its usability. The explanatory texts themselves offer good insight and are clearly worded. The design of the pages is modern and features a different color scheme compared to the rest of the app, with a dark blue background instead of a white one. The pages invite the user to find out more about the app and draw them in.

Login mask

On the login page you enter your login data to access the app. There is also a function for automatic login:
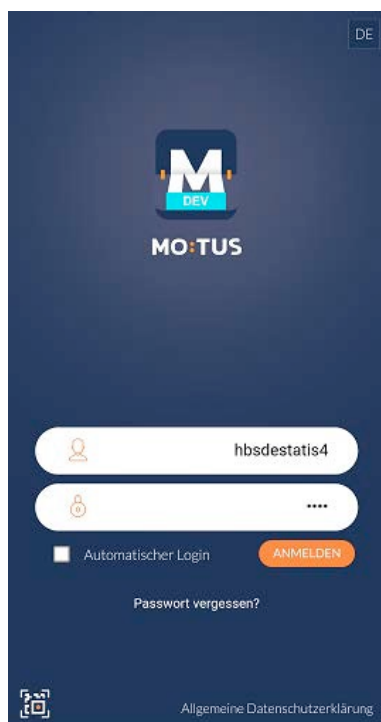
- Two remarks can be made about the login process:
  - On a smartphone with a curved display, the buttons are difficult to reach because they are located very far to the right. The exact location seems to vary depending on the smartphone size, for some they are too close to the corner.
  - Automatic login: the checkmark disappears even though the functionality works.
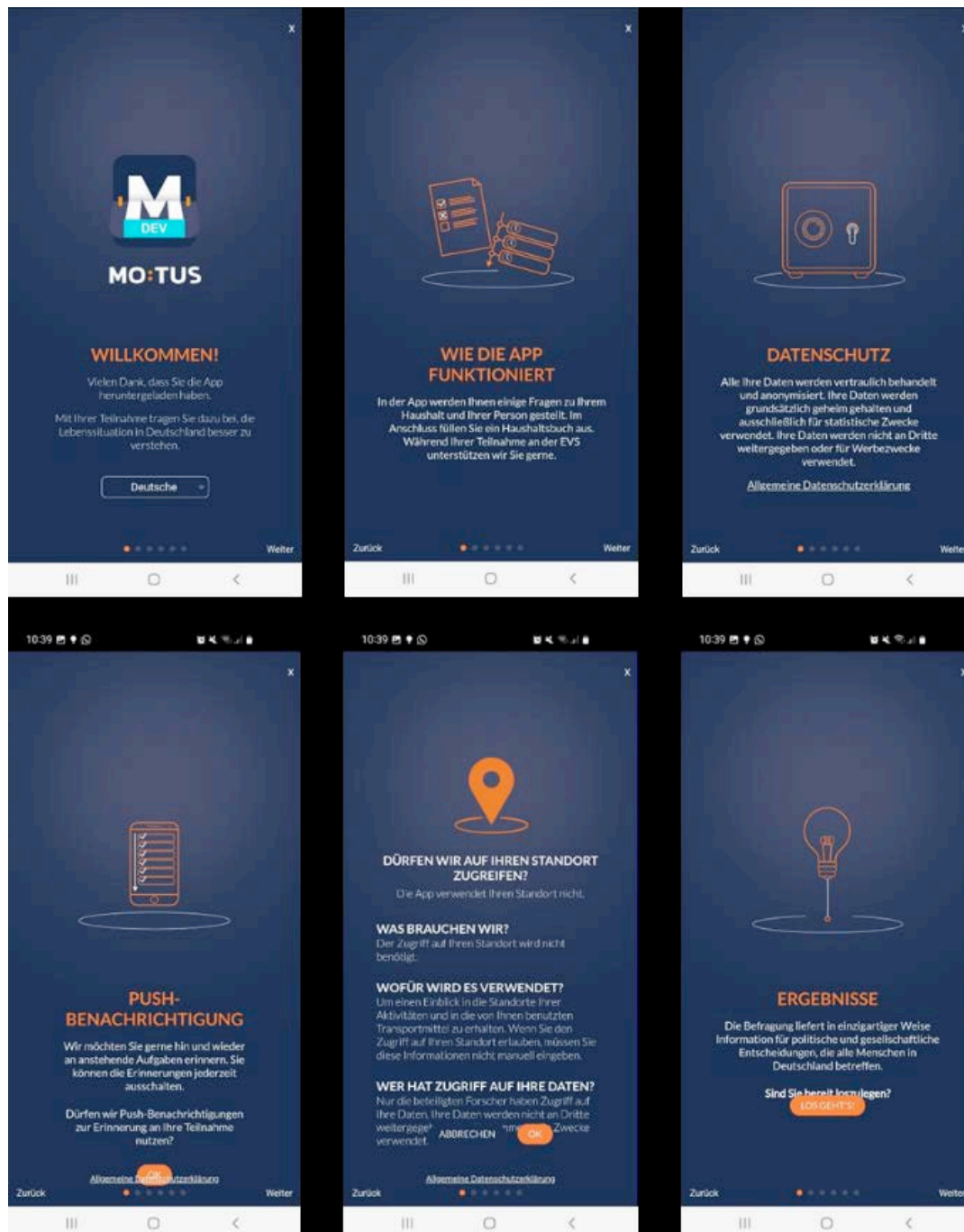
- The welcome pages give the user a little insight into the benefits and use of the app when they first start it. They also explain privacy, ask if push notifications are desired, and ask if user location data may be accessed.
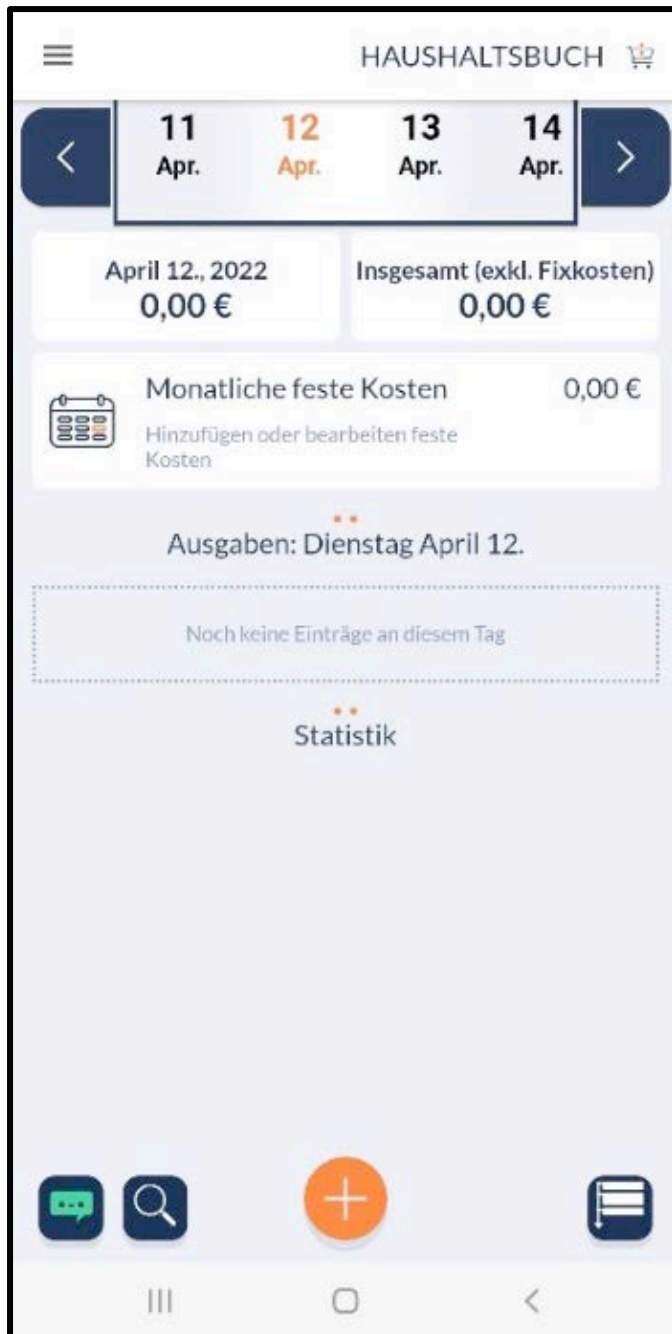- The design of the welcome pages was rated very positively as appealing and modern.
- The only critic our TP had was that it is not possible to access the page again. The users cannot read the instructions on the welcome pages again.
- Importantly though, several bugs were present in our test on these pages (see appendix for details).

Consumption Diary: Overview page

STATBEL
Belgium in figures

DESTATIS
wissen.nutzen.

VUB VRIJE UNIVERSITEIT BRUSSEL

h

eurostat

Capture new expenses

This page offers a calendar line for the selection of the date at the top, and below an overview of already recorded expenses per day and in total. Underneath is a larger tile to record fixed costs. At the bottom of the page, in the middle, is the plus button that can be used to add new expenses. To the left and right of this is the interaction bar, including icons for features such as the „assistant" or for a search (which, however, were not yet implemented during the test period).

- "Plus"-button:
  - All TP ultimately understood (mostly intuitively) that they had to tap the plus button to add new expenditures. However, this button was not found directly by all TP at first.
  - A TP first tapped the displayed square „No expenses for this day" in the expectation of being able to start input here. However, this field is inactive.
  - We therefore recommend to increase the size of the "plus"-button slightly.
- Tiles on the page:
  - Another TP first tapped the tiles. This TP complained that the function of the three tiles is not clear. The top two tiles provide an overview of the recorded costs and tapping causes a change of display to averaged numbers. However, the third tile opens a new page where the fixed costs can be entered. The appearance of the tiles does not indicate these different functions.
  - The advantage of this representation is (hopefully) that the fixed costs are not overlooked before you start recording the daily expenses.
  - We therefore suggest putting a visual indicator at the bottom of the (top two) tiles to show that you can interact here e.g. "Display daily expenditures".

Overview of expenses

- Generally, the overview of the purchases made (which is displayed after having added expenditures) has been praised. It was considered positive that one can see the expenses per day here.
- A TP would have liked to see the week's expenses at a glance in order to get an overview on which days she has entered something and on which she has not. At the time of testing, each day could only be displayed individually. Another TP also stated that she missed an overview where one can see for which days expenses have already been recorded. Without this overview, the function "no expenses" also has no perceived benefit for respondents.
- It is unclear how the amount shown in the right tile comes about. The value displayed is well below the fixed costs entered, although it is supposed to contain the fixed costs and the end of the month has almost been reached. Additional information is needed:

Subsequent processing of expenses

- If you tap on a recorded expense on the overview page, a blue bar with the option to edit this entry opens.
  - One TP was not aware that you can tap this tile. Two other TP criticize that it is not immediately clear that the option to edit the created purchase is hidden behind the pen symbol. To prevent this, the assistant should briefly mention the possibility to edit one's expenditures; a label could also be helpful.
  - The above mentioned blue bar was not immediately clear to everyone. A TP had issues assigning the blue bar to the correct expenditure (i.e. the expenditure above or below). We recommend to slightly adapt the blue bar so it can be more easily assigned to the correct expenditure by users.

Miscellaneous

Among the three symbols located at the bottom right, TP were not clear about the function behind the icon.

Please note: This was among the functionalities that did not work in the test version; however, as the icon was still mentioned as unintuitive (esp. in comparison to other icons), it should be considered replacing it and introducing it by the assistant.

One TP criticized the practice of stating that there were no expenses in a day. The plus button is probably interpreted by most as „add expenses", so the location of the option „no expenses" under the "add"-button made little sense for this TP.

One TP criticizes that the icon in the corner top right has no function, and recommends removing it:

New Expense or Return: Overview page

STATBEL
Belgium in figures

DESTATIS
wissen.nutzen.

VUB VRIJE UNIVERSITEIT BRUSSEL

h

eurostat

- This page serves as an overview of the to-be-added expenditures. Initially, only three tabs are visible: date, country of purchase, and shop. Afterwards, more input fields are shown, and this page serves as the overview for all single products bought at a single shopping trip.
- Generally, some TP claimed that only when seeing the categories for the first time, they realized how many different expenditure types are to be put in. They wished to have been sensitized about the variety of possible entries earlier. Not everyone thought of also entering non-food purchases e.g. in the hardware store.
- When ordering goods from other countries online, it was not clear which location had to be chosen for the entry. Below the entry of the country, it was asked where the shopping took place. This irritated TP, as the country was already chosen, but "where" was now meant as the name of the shop. Further having the sum of zero at the bottom of the page was not considered necessary and was perceived as rather irritating.
- In case of a day without expenses, TP did not realise they had to confirm that in the app via the button on "new expense or return". Rather, they would just not type anything until the day the next expense was made. Visual cues and push-notifications could help as a reminder after several days without entries.
- For the German version of the app, the date-format should be changed to the languagetypical order "weekday, day, month".
- For every new expenditure, "Belgium" was the default setting and thus had to be changed with every new shop entry.

Shop search
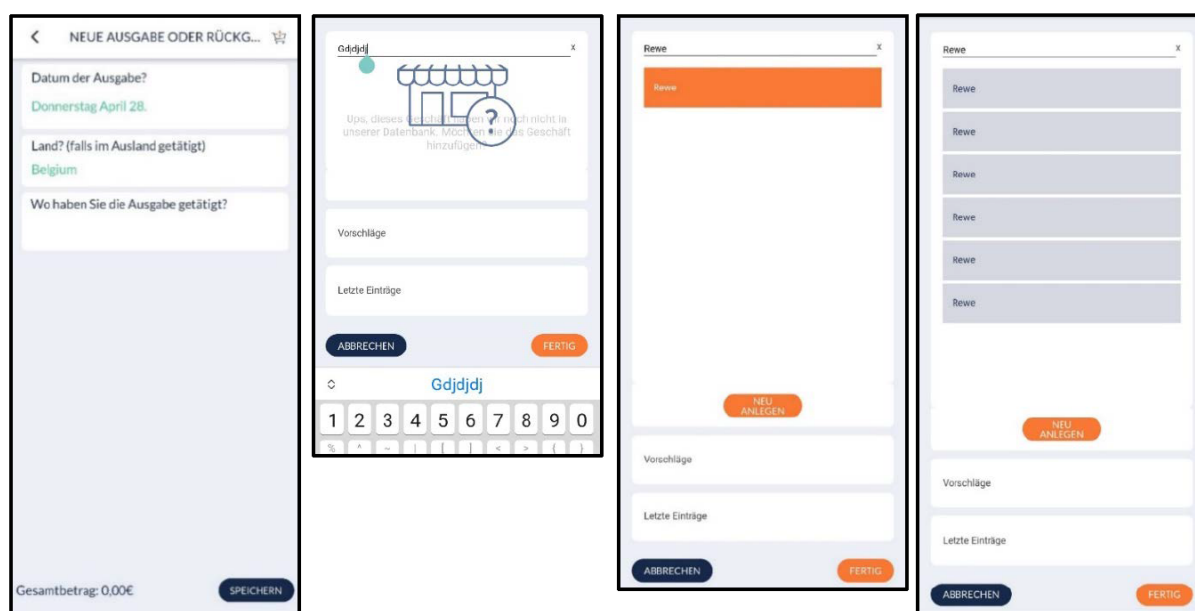
- Adding a shop is done by clicking on the third tile of the initial overview page, which opens a search window.
- Having set the country, users get to enter the name of the shop, by default via a search. Either, the shop is already listed and appeared in the search result list, or it had to be added manually to the list of shops (please note that the functions "suggestions" (Vorschläge) and "last entry" (Letzte Einträge) did not function). Only when one suggestion (orange) is actively chosen, the entry was saved after clicking on "finish".

- Some participants, though, did not realize they had to press the "finish"- button after having chosen a shop. They clicked on "add" instead and got to the option of matching the typed shop to the respective category, even though that shop was already listed in the app.
- Other TP tapped on "ready" without tapping the search hit in the list. They wondered why everything is deleted upon clicking the button, and they had to start the search again.
- In case a shop was not found, instructions for users should be more elaborate. Our suggestion: In order to add a new shop, the button "add" ("neu anlagen") could be named more precisely, e.g. "No matches? Add a new shop here".
- Additionally, the formatting of this button should slightly differ from the "finish"- button (less salient), so the user's attention is not automatically drawn towards adding a new shop.
- In terms of design some TP had issues when the keyboard was open-ended, as the image of a shop and a question mark were overlapping the grey text, which is therefore difficult to read. Also, in this state, the "add"-button was not visible any longer. Only when the keyboard gets closed does the "add"-button re-appear for these users (issue when the smartphone display is not large).
- When participants wanted to add a new shop to the app, they could link it to a category (like supermarket, service, etc.) and afterwards a sub-category. However, once having reached the subcategories, there was no option to go back to the previous category. Currently there is no "back"-button, only a "cancel"-button, which stops the entire process of adding a new store and puts the user back to the shop search. Using the phone's "back"-button also restarted the whole entry process. We therefore recommend to change the functionality of the button and rename it to "back".
- When choosing a shop category, there is some overlap among them. TP report having no clear definition about which category to choose (e.g., when it's an online purchase from a retail store that has several offline branches). Along these lines, when choosing a shop from the list, it currently is automatically depicted as a regular offline store, even though nowadays many stores offer an online purchase service, e.g. supermarkets provide the option to purchase online and deliver the groceries to one's home.
- For some TP it was unclear whether the shops that are listed when searching are already linked to a shop category or whether they have to be added manually (but did not know where). Only after going back to the overview page, one can see that a category is added automatically.

Product Search

Overall Evaluation of Usability and Functionality

The product search, as one of the central and most used features of the app, has been evaluated in depth. We had our TP use the functionality to enter their daily expenditures over the course of one week, explored the topic in our focus group, and also conducted internal functionality tests with a variety of different mobile devices (including different brands, models, and operating systems). This allowed to us extensively test both the usability and the functionality of the product search.

- Overall, the product search works well – if a product is listed in the database and a user types the exact name into the search, the result is (usually) correctly displayed at the very top.

- However, several issues currently diminish its usability:
  - In some instances, even an exact match is not displayed prominently at the top.
  - Even very small variations can strongly impact the search result, leading to no (relevant) items in the results.
  - Furthermore, while groceries from a supermarket are mostly included, the search lacks several products or services from other expenditures (such as car repair, tickets for cultural events, et cetera).
- The functionality tested quite well, the app display is coherent and comparable over several different devices, brands, and browsers. We found only a few, mostly uncritical bugs.
- Selection of screenshots showing the product search:
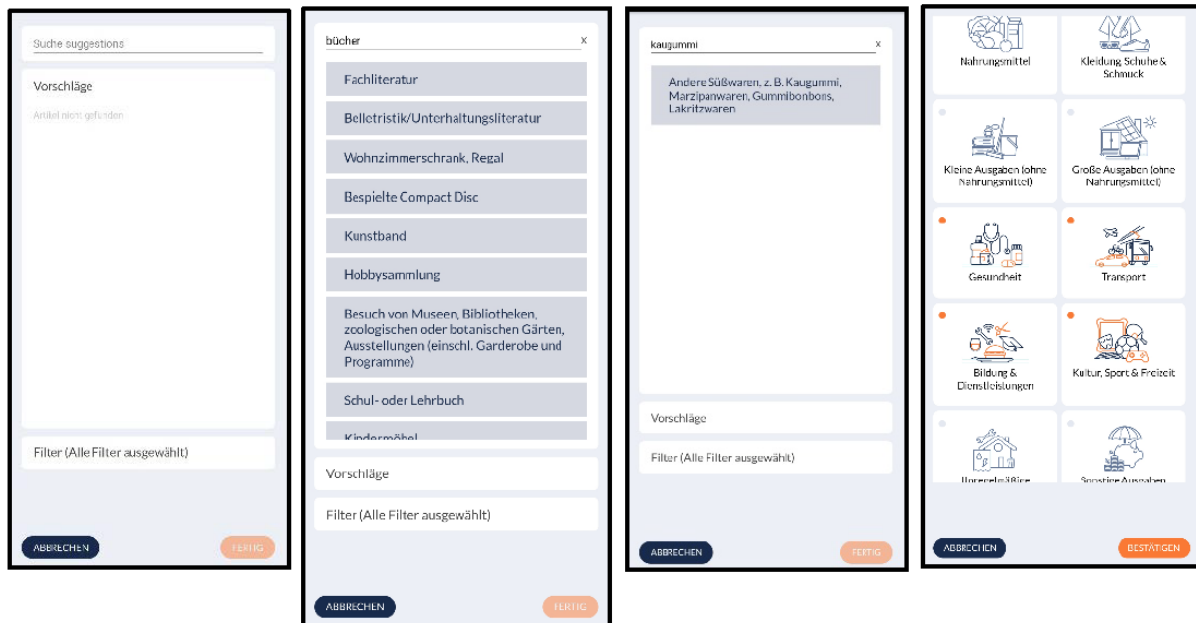
Concept

- We recommend including more assistance, to provide basic information for respondents on how expenditures are supposed to be entered into the app.
  - For the TP it was unclear at first, if and how they could enter their e.g. supermarket expenditure in one single entry ("supermarket xyz: 60€"). They lacked understanding of the app concept, i.e. having to add each product manually. This might be more prominent in Germany, where the HBS is conducted over the period of three whole months (and for the most part, only the total amounts have to be registered).
  - The handling of already prepared meals, especially how to add restaurant visits (including reported search terms such as "lunch", or "canteen" or "food delivery") posed a problem for several TP. They did not find any matches, and were lost how to enter such expenditures into the app. While some realized that this is (partly) supposed to be entered as the place of purchase instead of the actual product, they still oftentimes did not know what item to enter – How to add a full meal consisting of several products, such as a steak with sauce and vegetables?
  - To avoid this, we recommend providing more assistance on what is supposed to be entered where and on which level of detail (for example, via a tutorial when a user first opens the search).

- Along these lines, TP struggled to enter expenditures whenever the search did not yield proper results (see chapter 'Search algorithm' for details on the issues TP had). This led to omitting expenses for which they did not find fitting search results in some cases.
  - This issue was the more prevalent the further away the product was from the typical "supermarket" product, e.g. car repair, conference fees, entry tickets, …
  - We recommend to add a button that allows for a manual entry (and optionally, if feasible, addition to the search list) of an expense, if the search does not yield any (fitting) results. Ideally, clicking the button would not simply open a text bar, but instead ask the respondent to go top-down from general expenditure categories to their actual product/service bought. This would allow easier manual coding to the correct COICOP in the following data processing and editing stage.
- A per se helpful, but in its current form quite cumbersome feature of the search bar is its automatic search, which triggers after roughly one second of inactivity when typing. While this potentially saves users an additional click on "search" (and for some even the need to close the keyboard in an additional step first), the period was considered too short by many TP. Triggering the automatic search closes the keyboard and starts a quick search. If respondents have not actually finished their search term by then, they have to re-open the search by clicking into the search field again.
  - If search terms include special characters, umlauts, etc., the short period makes it near impossible to enter such words in one go on a smartphone. This is because on a smartphone keyboard, you have to keep a letter pressed for a short duration to get the options of adding these special characters (e.g. to get from "a" to "ä"). Most often, it triggers the automatic search first before being able to click the special character.
  - Over the period of one test week, adding all expenditures, this turned out to happen fairly often and many TP were irritated by this.
  - Furthermore, it causes a form of stress for respondents when adding expenditures, as you want to type fast enough to avoid triggering the search prematurely.
  - → We recommend increasing the period until the automatic search activates by one or two seconds, as the feature in itself is handy for respondents.

Layout

Search bar

- The search field should be slightly more visible. Currently, it is rather small and located inconspicuously at the very top, causing respondents to initially overlook it and first point their attention to the less important options "suggestions" and "filters".
  - Furthermore, the "Search" text, titling the central feature on the page, is in a smaller font than the options "suggestions" and "filters".
  - We therefore recommend to design the search field, as the central interaction point, more visibly (e.g. slightly larger bar, or more centrally placed with an automatic shift to the top area when a user starts typing) and increase the text font of "search".

- The option to clear the current search term via the "X" icon on the top right is very small, which makes it barely noticeable and hard to tap. We recommend increasing its size slightly.

Display of search results

- At present, if a search yields more results than fit the smartphone screen and the list is cut off right after the last visible search hit, there is no visual cue indicating the possibility to scroll down for further results. This can lead to omitting results by users. We thus recommend adding a small scrollbar (or a similar unobtrusive marker) to the right when scrolling is required.
- As an idea to promote the visibility of exact matches which might not (all) be displayed at the very top, either due to several results including the search term or the way the chosen search algorithm functions: It could be helpful to slightly boldface the exact search term matches, e.g.:

- In rare instances of especially long text, it currently gets cut off. While is this probably hard to fix from the programming side, experts responsible for inputting search results should be aware of a maximum character limit, and if exceeded, provide fitting breakpoints (using "-"). An example of this can be seen in the screenshot above (second result).
- A minor issue, but to further increase the user-friendliness, it should be possible to deselect a previously selected category. Currently this is not possible, one has to pick a different category or alter the search term.
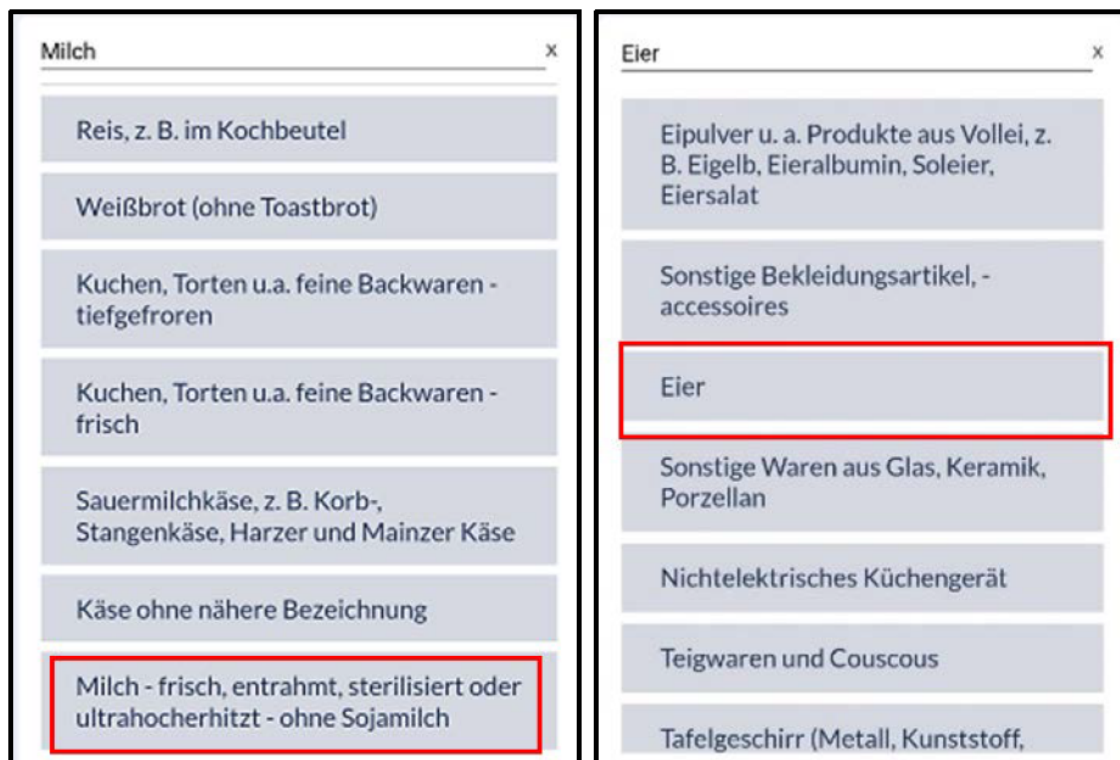
Algorithm and item list

The search algorithm and item list used for the test worked acceptably for most searches.

Nonetheless we could identify several key points of improvement that should either be incorporated into the current algorithm, or be considered as input for choosing a different algorithm.

General issues

- In some instances, categories containing an exact search match are not displayed at the very top, but instead appear somewhere further down the list, after less relevant results. For users it is not comprehensible why a perfect match is not listed as the first result. We therefore recommend to list results that incorporate the exact search term at the top.
  - Examples "Milch" (milk) or "Eier" (eggs): The first few results are (partly) in no direct relation to the search term, while another entry contains the exact match.

STATBEL
Belgium in figures

DESTATIS
wissen.nutzen.

VUB VRIJE UNIVERSITEIT BRUSSEL

h

eurostat

- For some products, the exact name does not yield any search results, even though the broader category to which it belongs exists and lists this exact product type as an example. As this is not comprehensible and therefore frustrating for users, any example explicitly mentioned within a broader category should also be factored in when searching for this exact product.
  - Example: Butter cheese ("Butterkäse") vs. Cheese ("Käse"): The term "butterkäse" does not yield any results, even though it is specifically mentioned as an example of the broader search result of "Käse":
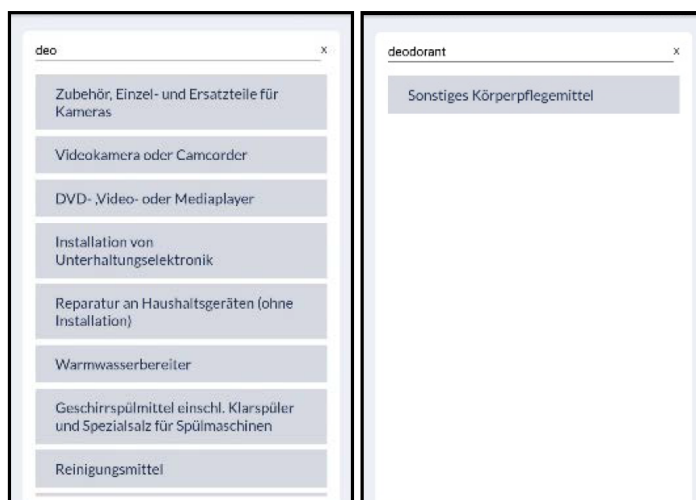
- Generally, if the exact search term exists within a category (even if not the full name but only parts of a search result), it should be 1. among the search results and 2., depending on how many characters fit exactly, displayed further up the top.
  - This would help reduce the already mentioned issues with colloquial terms that are abbreviated in everyday life, such as "Deo" for "Deodorant" in Germany, or "Zug" standing for "Zugticket" (in both examples, the search for the short version did not yield any results, while the formal complete word is among the available search terms).
- The search (partly) lacks more general terms (tags and/or higher COICOP labels). These should be added and linked to all the available detailed products on the corresponding lowest COICOP level (for which products are supposed to be entered). To illustrate what is meant, our TP had trouble adding simple items such as "Shoes" – searching for it currently only lists two results, both rather irrelevant. However, the list of possible search results contains detailed types of shoes, such as "Sneakers". Therefore, we recommend to list these detailed types of products when a broader product category is being searched. (Note: For some categories, this is already implemented, but overall seems fragmented and should be expanded upon).
  - Example: "Schuhe" (shoes) leads to no proper result, while "Sneakers" lists the right category "Straßenschuhe" (street shoes):

- The search list furthermore lacks some items in colloquial or abbreviated spelling which should be added as tags linking to the corresponding "official" item label.
  - An example of this is "Deo" (showing a lot of unrelated results), short form for "Deodorant" (showing one fitting broader category where it belongs to):

STATBEL
Belgium in figures

D STATIS
wissen.nutzen.

VUB VRIJE
UNIVERSITEIT
BRUSSEL

h

eurostat

- Singular or plural variants greatly impact the search quality. Ideally, it should make no difference whether one searches for books ("Bücher" – leading to mostly fitting search results) or book ("Buch" – mostly unfitting categories).
- Topic of already prepared/premade food: This turned out to be an issue for TP, as they did not know what to search for when they had a meal at a restaurant or delivered to their homes.
  - Neither general terms, such as "meal", "fast food", or "lieferando", nor the specific dish's name, e.g. "Kebab" or "Ramen", lead to satisfying results.
  - It also leads to uncertainty as to what should be added into the diary – try to find the meal as a whole / its dish name, or add the main ingredient(s)?
  - We thus recommend to add the more common dishes, and supply these (among other results) when more general terms are being searched (such as "meal").
- It is very positive that the search list already contains several brands (which are often used synonymously with the actual product name in everyday language).
  - Ideally, the list can be expanded even further with additional brands (or rather, their 'main' product/s) by sifting through all product categories and adding the 1-2 most common brands/product names respectively. Examples of non-included brands are Bepanthen (category "skin cream") or Zewa (category "paper towels").
- To further enhance the user experience, the search results list should be constructed in a way that ideally also favors the most common products and puts them in the first place. Searching for flour ("Mehl") lists "other types of flour" as the first result, which appears more like a miscellaneous "other"-category than one that should appear first. Currently, after this result, several unrelated items follow, and then somewhere further down the list, the most common type of flour, wheat flour ("Weizenmehl"), is listed. With the improvements already mentioned, the unrelated items should be gone or moved further to the bottom of the results list, but ideally, wheat flour should be the first result here, followed by the second most common type of flour, and then at some sensible point, the "other"-category.

Specific issues

- Spelling errors immediately lead to no results. To increase the user friendliness, the search should account for typing errors and still provide most probable results.
- Adding a blank space influences the search results. This is especially problematic, as smartphone keyboards oftentimes offer suggestions when typing, and for our TP, choosing one of these suggestions automatically adds a space after the word. (Also currently contains a bug: not only does it not provide helpful results, but in its current form, it always leads to the display of the same four search results, completely independent of the actual term searched). The search algorithm should ignore blank spaces if possible (or at least when blank spaces are not followed by additional text, i.e.: "Käse_" treated as "Käse") o Search term without (left – correct) vs. with a blank space (right – bugged):

- Umlauts (ä,ö, …) and accents (á, ó, …) should automatically be incorporated in a search, without the user having to type them out in the exact same way, i.e. such characters should be treated just like basic characters (a, o, …), and in the case of umlauts additionally as "ae, oe, …". As of now, umlauts are not recognized. For example, typing "Kaese" instead of "Käse" leads to no results. This is especially problematic, as some receipts generally do not use umlauts, listing it as "Kaese". This increases the probability of such spelling, as it prompts users who enter their receipts manually to type in an items' name one-to-one.
    - Example: Search with "ae" instead of "ä" leads to no results:

- Various special characters (e.g. #) within a search term cause the app to crash. This should be fixed, as some brands or product names might contain such special characters, or they could be added by users unintentionally in the form of typos (see bug list in the appendix for details on which characters were identified as problematic).

Additional options "Suggestions" and "Filters"

Unfortunately, in our test version both the Suggestions and Filter options did not function (see appendix with Bugs for details). However, we could still derive a few theoretical tips to strengthen its usability.

Suggestions functionality

- As the suggestions functionality did unfortunately not work, it remained unclear how exactly it is supposed to work in its current implementation, i.e. whether it serves as a list of recently added products or actually tries to suggest fitting (new) products based on expenditures entered thus far.
- Generally, it would be useful to have a list of all already entered products, as it is likely that the same product gets purchased twice over the course of the diary duration (especially in Germany, where it lasts for three months).
- Ideally, the list can be sorted manually by either topicality (products entered most recently are displayed at the top) or frequency of selection (products that have been entered most frequently are displayed at the top).
- If the functionality is indeed meant to display previously entered articles, we recommend to rename it, to communicate its functionality more clearly, i.e. rename it from "Suggestions" to "Recently added items".
- Moreover, as long as no entry is available, a short but informative text should be displayed there (e.g. "No recently added purchases available") rather than the current, generic text "No article found".
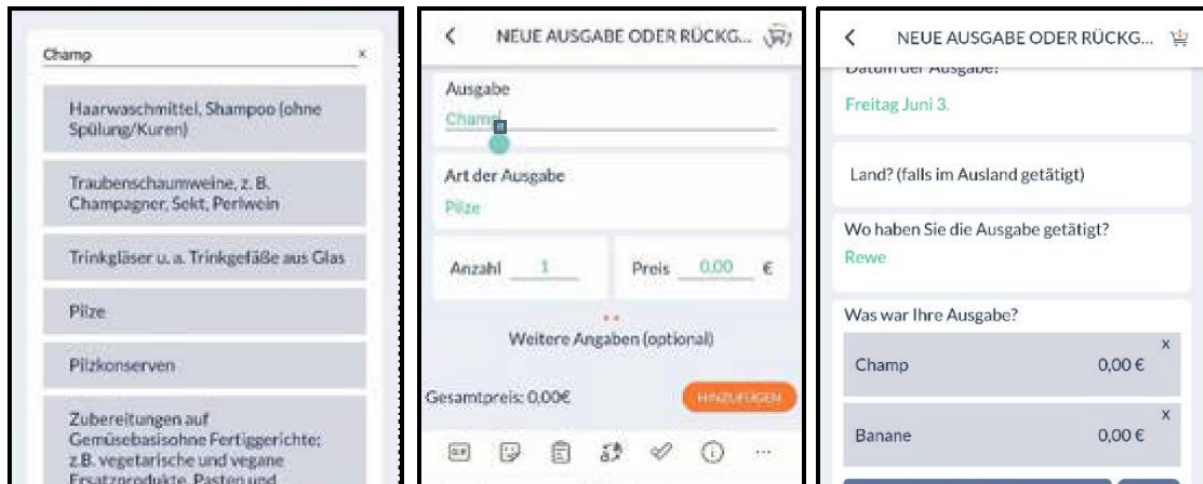
Filter functionality

- The filter functionality is in theory also a helpful tool to refine the product search, in the context of the thousands of products available.
- To improve its usability further, it should complement the frequent use case of only searching within one category, e.g. groceries/food.
  - If a user can only think of or only has a short, more generic search term, the search might yield a long list of possible expenditures. To narrow it down, disabling all but the fitting category is the easiest way to handle this problem. However, currently, one has to manually disable every single category, and only leave the fitting category ticked to achieve this.
  - To increase its usability, we therefore recommend adding an option to disable all categories, allowing users to search within only one category by ticking solely the one single adequate category, thus saving them several clicks.
  - Alternatively, an option could be to disable all but the one category a user has clicked on by default. As this is likely the most frequent use case.
- Within the recommended addition of an assistant, an info should be added that explains the filter functionality and its relation to the search results.

Adding details to an Expense or Return

General remarks

- Adding every single item bought was perceived as burdensome. Once having typed in several products, people might get used to solely entering the product-name and price. In that same vein, TP asked for a function to enter the sum of the receipt as one, e.g. "Rewe, 52€". This would be a large reduction of respondent burden, since it concerns a timespan of three months in Germany.
- Another issue arose concerning the order of entered items.
  - TP copied products of supermarket-receipts chronologically into their phone. The app, though, shows the most recent entry on top of the list, which means items are in reversed order compared to the original receipt.
  - In case the user gets less attentive while typing and is uncertain about the product entered last, he has to scroll up the list to check for the most recent item, as the item added most recently is not depicted as the one next to the add button, but at the other end of the list instead.
  - Reversing the order of items in the app would help users, especially for a final check whether all items were typed in.
- When entering products manually in the search field, suggestions popped up before the word was typed entirely. TP then chose the matching suggestion, expecting the word to be auto-completed (as is done in many search engines). If not completed in the next step, when details of the purchase had to be entered, the half-finished word appeared in the list of products. Some TP criticized this and preferred auto-completion.
  - However, a possible auto-completion of words shows pros and cons: The advantage is that TP are less irritated by incomplete item-names in the list. A disadvantage is that classification sometimes do not match the wording of products which is used in spoken language. Our following example shows an entry- try of "Champignons":

STATBEL
Belgium in figures

D STATIS
wissen.nutzen.

VUB VRIJE UNIVERSITEIT BRUSSEL

h

eurostat

- When entering a lot of products, it gets harder to keep an overview, as one needs to scroll back and forth to see all items. In order to increase visibility, line height could be diminished, so that one can see more products on one site without having to move on the page. Usually, people are used to the compact view of items on receipts, so one can display items tighter (ideally done via automatic rescaling depending on the amount of items entered).
- Another design remark was made by one test person. The rather large (blue) button "add a new expenditure/return" could be replaced by the orange "plus" button used when entering shops. This is because users already learned the meaning of the orange "plus". No need arises then to familiarize oneself with another button. The current blue button does also not visually indicate that one has to tap it to add products one by one.
- Items typed into the app can easily be deleted. One TP questioned why one should even use the return function, if deletion is so much easier. An instruction could be added, which explains why the use of the return function is important. Alternatively, when clicking on the "x" to delete the light-blue tile with an item, a warning could pop-up saying "Are you sure to delete? If you want to return the item, do the entry here", with a link to the return function. Optionally it can be un/ticked by users whether this warning should be displayed again in the future or not.
- When entering cent-amounts, a 0 before the decimals is mandatory to type into the app. An "automatic 0" that appears when typing a decimal separator like dot or comma would save users a click and therefore be more user-friendly.

Item counts and measurements

- TP tended to mix up the entry of the following two types of purchases:
  - 1) having bought two or more pieces of the same item, like five cups of cream or three packs of butter and
  - 2) goods consisting of several entities, like a package of apples or oranges.
  - The number of items ("Anzahl" on the screenshot) referred to the first case, where the same product was bought several times, while the second option referred to 2):

- The price next to 1) ("Anzahl") gets multiplied by the number of items entered, which leads to a new total price. Everything entered below the two dots in the middle does not result in changes of the price and is therefore voluntary. This applied for the second case described.
- Often people tried to enter a pack of fruit (e.g. mandarins). They put 6 for "Anzahl" and then the price for the whole net, e.g. 2 Euros. When 12 Euros appeared for the total price, they noticed that obviously something must have been done wrong. At that point, TP understood that the number of pieces inside a pack had to be given below as an optional entry.
- TP wanted to fill in everything as thorough as possible. They did not notice the label "optional" and did not know that information was voluntarily. For them, it felt mandatory to fill the fields out. A thorough communication of voluntary input as well as more precise instructions on data required could help users reduce the burden.
- We recommend to change the structure of the page as follows:
  - Price is one of the most important tiles. It should thus be aligned to the left to strike ones' eyes and increase its visibility. Prefilled numbers should be avoided.
  - The "add" (Hinzufügen) button is bright and visible and even clickable if no price entry was made. We recommend to keep this button inactive as long as no price was typed in. Alternatively, an error message should be added if the price is missing.
- Concerning the use of units, the following issues arose:

  - The field "measured in" was perceived as confusing, as its label (unit of measure, e.g. "g", "kg",…) is rather far from the number field.
  - Having chosen to enter an item in kg, the app did not allow values below 1. No instructions were given either as to which unit should preferably be used.
  - Gramm should be used as a standard unit. Rephrasing instructions concerning the desired and optional ways of data entry could lead to more certainty among users. Typing decimals in kg could automatically be transferred into grams, as this saves one click.

# Returns: Discounts, deposits, and other types

- After the one-week test phase, two TP were still not aware of the fact that returns, deposits etc. were also meant to be added into the app. Additionally, some TP mentioned that they perceive it as too burdening to enter these as well.
- One TP wished to enter the discount in percent and not in euro. In case of a percental discount, the respective amount had to be calculated manually. The app did not offer a possibility for that.
- As receipts often display the price of discounted goods next to the names of the items, TP saw a tendency to just copy that price. In the end, when being asked for discounts, they additionally copy the sum of savings to the discounts, which is often outlined at the bottom of the receipt. In this case, discounts were subtracted twice.
- Further, if discounts are only displayed at the bottom of the original receipt, users might put it in lastly. If they only see the top of the page on the smartphone, though, they need to scroll down in order to enter the discount sum. Therefore, it might be forgotten by users.
- We recommend a more thorough guidance to the user at that point, especially in terms of choosing the right type of discount. It should might be worth reconsidering, if the function is really needed on product level or enough on the level of the whole purchase. Otherwise, people might add discounts on product level manually or add the sum of savings on the bottom of the receipt.

- If discounts are added for the whole purchase, one can choose from the categories on the left. However, we found these four categories to not be exhaustive. A category for bonus programs such as Payback, as well as an "other"option could be added.
- The third option is translated as warranty ("Garantie"). It seems unusual to have this included as a type of discount.
- TP suggested to put the category deposits here, as it was mentioned on the button they clicked on one step before.
- TP would also like to have quantity discount (on item-level) and discount on abonnements (Netflix, etc) to choose from.
- As there are two direct ways to enter discounts into the app (as a detail of an item and based on the whole purchase), the question of the detail of discount arose, as it is only being asked on the level of the whole purchase.
- Deposit for bottles is listed among the description on the overview page, however, the actual option when clicking on it (screen to the left) is not available.

Fixed costs

The monthly fixed costs are entered via a tile on the overview page, which opens a new window with 10 categories and further subcategories. (Note: at the time of the test, these were still in English):
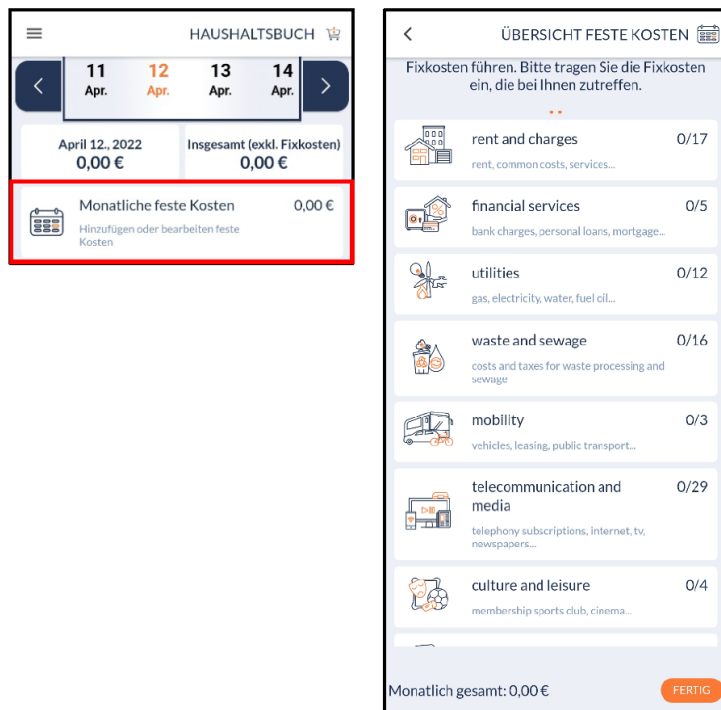
Concept

- When it comes to fixed costs, some TP asked for more information about how they are defined. We therefore recommend inserting an info-"i"-button that contains explanations.
- The display of the amount of subcategories (on the far right of every list entry, e.g. "0/17") was confusing for some TP. As the display inherits not much benefit, we recommend removing it.
- If you tap on a category, further subcategories are displayed. One TP criticizes that there is no „back" button here, only „done". We recommend adding a "back"-button on the left bottom of the screen. This makes it clear how to leave the page without entering anything.
- A search function would be helpful for categories with many subcategories (e.g. insurances).
- When entering a fixed cost, there was the following problem: after specifying the expense and specifying the amount, you have to tap „add" twice and only then press „done" to conclude the addition. However, the „add" button is often being overlapped by the keyboard (unless the smartphone display is rather large). So, after entering the name and amount, TP tap the „done" button. However, this means that all information is deleted and you end up on the previous page (main categories), since without clicking on "Add" first, no entry gets saved. This was very frustrating for some TP. We therefore recommend to disable the display of the "done" button, i.e. not fixate it, but let the keyboard overlap it:

- When specifying the periodicity, the option „quarterly" is missing:

Categories

- Concerning the content of (sub-) categories, TP missed the possibility to add public transport subscriptions. Further, one TP pointed out that there is an overlap between the categories „online streaming service subscription" and „netflix".
- Underscores are visible on one entry: „bank_account_costs":

Options menu

- Generally, several of the options within the burger menu were not properly set up for our test. However, we can still give a few recommendations for improvement to specific points.
- Within the options menu (accessible via the burger menu), there is a profile page where you can change settings for contact information, preferred language, time zone, password and username.
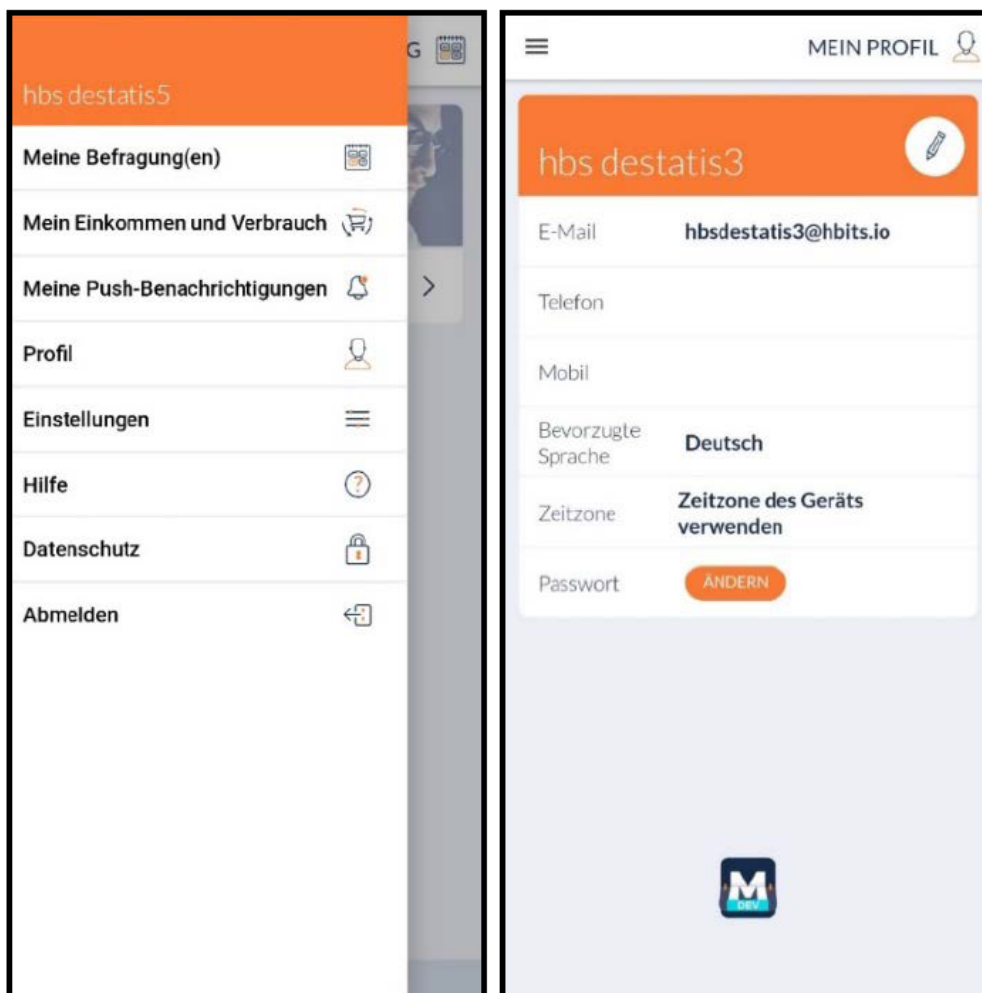
- o The pencil icon on the very top has to be tapped in order to change profile data. On the bottom a "Change" button is displayed prominently, however, that only applies to the password. A few TP criticized that they did not understand at first how to change their data. The pencil icon was not recognized, and they clicked on the bigger "change" button instead, which only allowed them to edit their password. We therefore recommend to emphasize the pencil icon more and add a label to it. o The button to change the password should be renamed to "Change password" to communicate its functionality more clearly.
- o The language selection option should be located in the options menu instead of the profile page, as users will intuitively search for it there. In addition, the options for geo tracking and notifications did not work yet.

„MyGroup"-Functionality

About the MyGroup functionality:

- The MyGroup functionality supports the creation of a household and subsequently the addition of household-wide expenditures, i.e. splitting or copying expenditures within the household.
- Once the assigned head of a household provides the number of household members, new identifiers are created and every member receives an individual questionnaire. Once the questionnaire is completed, everyone is held in a synchronization phase until all household members have completed the individual questionnaire. All participants who have reached the synchronization phase can view the current status of the remaining household members. Once everyone has completed the individual questionnaire, only the head of household can send in all household members and the rest of the survey is unlocked (e.g. the main diary).
- At the time of this test, the rest of the survey was not available, and some of the process until there was preliminary and / or only rudimentarily implemented.
- Please note that the functionality was also only programmed for the WebApp, not the native mobile app.

Functionality test results:

- A functionality test was done to test the current implementation for technical issues; TP did not see this feature.
- As the test possibilities were very limited, only a few basic functionalities could be evaluated.
  - The features that were available all functioned without problems (ignoring issues in the individual questionnaires that were not properly set up in the first place.)
  - Both, the confirmation mails to the household head as well as various mails to the household members, were sent correctly.
  - After having completed the first questionnaire, a popup that notifies the respondents about successfully finishing it worked as well.
  - The synchronization phase also did not show technical issues and worked well in its current state, both for the head as well as the members of a household.
- A positive result was generally, that there were no significant differences between the Google Chrome, Mozilla Firefox, Microsoft Edge and Safari browsers.

# Addendum – Bugs & Technical issues

In the following addendum, all technical issues and bugs are listed. They are roughly sorted by the main chapters of the report. If no bug was identified within a chapter, it is not listed.

Login & Welcome pages

- The orange dot that should indicate the progress remains unchanged when advancing through the pages:

- Several design bugs with buttons overlapping texts:
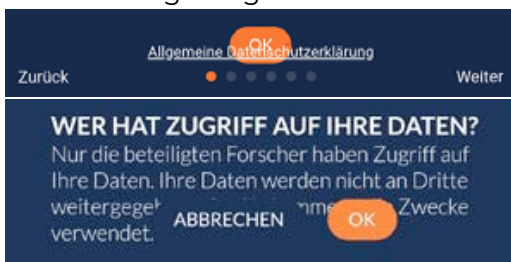
- On the last page, the continue ("weiter")-button is still visible, but does not work. It can be removed, as there is a larger main button "Los geht's" that is meant to be interacted with:

- The selection of the location leads to a crash of the app. The app loads endlessly and only a restart can fix the problem. This happened independently of the smartphone brand:
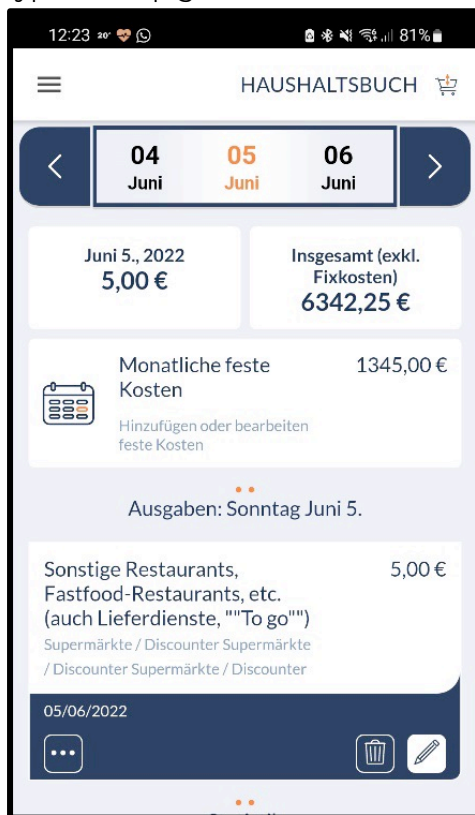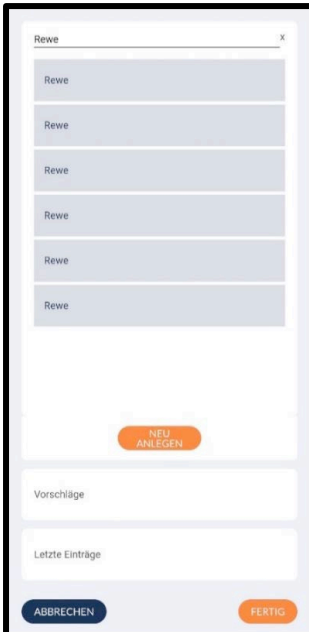
Consumption Diary

- Subsequent processing of expenses: The function behind the three dots was inactive.

- Using the function to edit expenses automatically changes the specified country from „Germany" to „Belgium".
- Display of expenditures added: When one changes the location retrospectively, the type of shop gets listed several times:
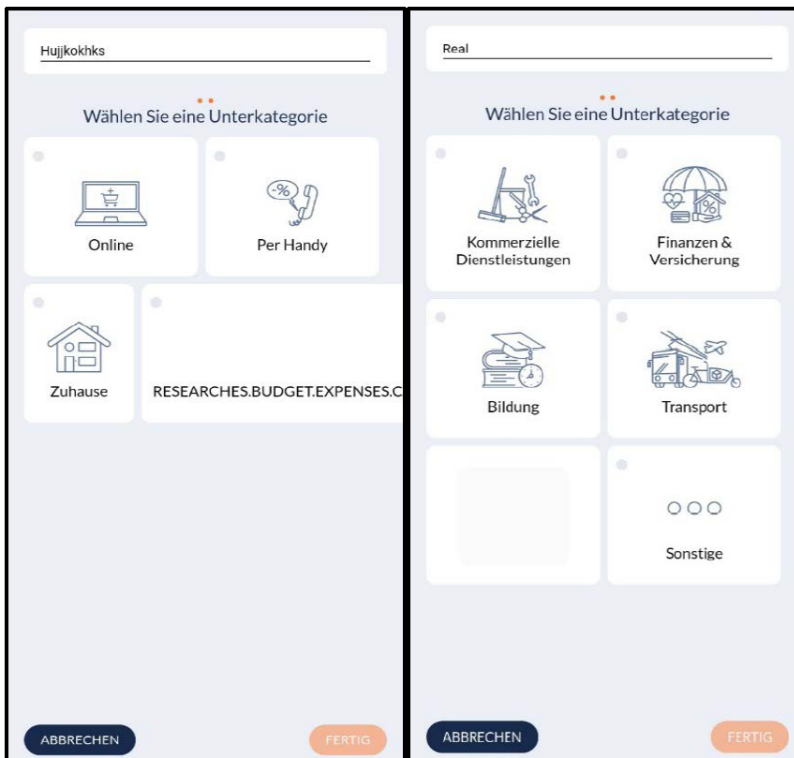
New Expense or Return: Overview page

- Repetition of shop-names: When entering the correct name of the shop, its name appeared repeatedly according to the number of times it was used before.
- Having typed in a number and then deleted it, the space for the cursor disappears and the underline for typing the number jumps up right below the naming of the field.


- Some shop categories showed bugs in their subcategory selections (empty tiles and/or internal terms/codes):

Product search

- Two TP experienced a crash of the app when typing a product name in the product search. They got both logged out and landed on the Login page.
    - Devices this (among others) happened on were: Samsung Galaxy S 10 w/ Android 12 & Xiaomi M11 w/ Android 11 (MIUI Global 12.5.7)
- When typing something into the search list, and then clicking on the arrow- or checkmarkicon on the smartphone keyboard on the bottom left, the app crashes and you land on the Login page.
    - Devices where this (among others) happened on were: Samsung S21 5G w/ Android 12 & Xiaomi M11 w/ Android 11 (MIUI Global 12.5.7)
- Special characters within the search cause the app to crash:
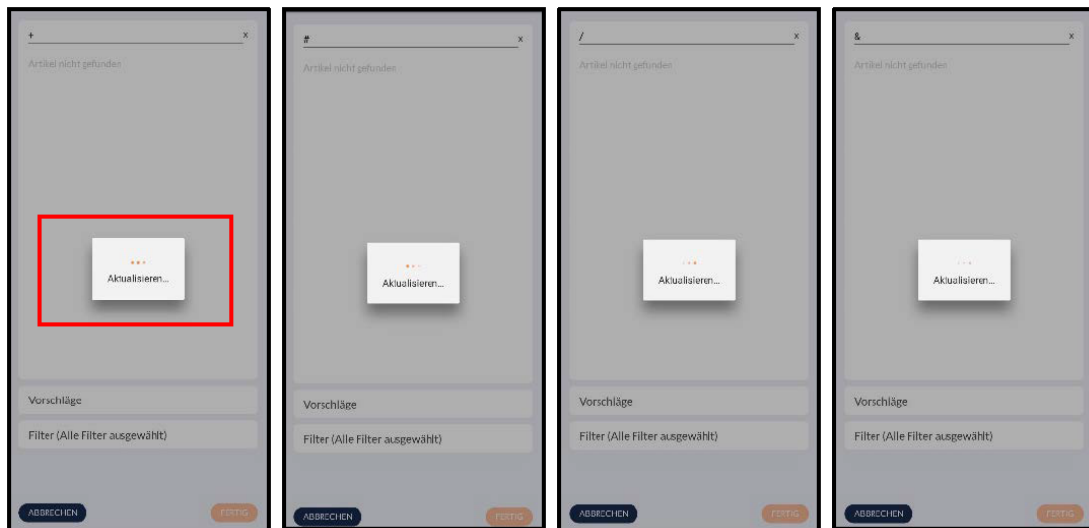
    - → Using various special characters within a search term cause the app to crash. After typing these characters, a permanently visible "Aktualisieren" (updating) overlay is displayed, which cannot be closed manually (pressing the backwards-button on the smartphone does not work either). One has to completely shut down the app and restart it.
    - This bug occurs with the following special characters: "+", "#", "/", "&"
- The inclusion of blank spaces currently leads to a bug:
  Search term without (left – correct) vs. with a blank space (right – bugged):
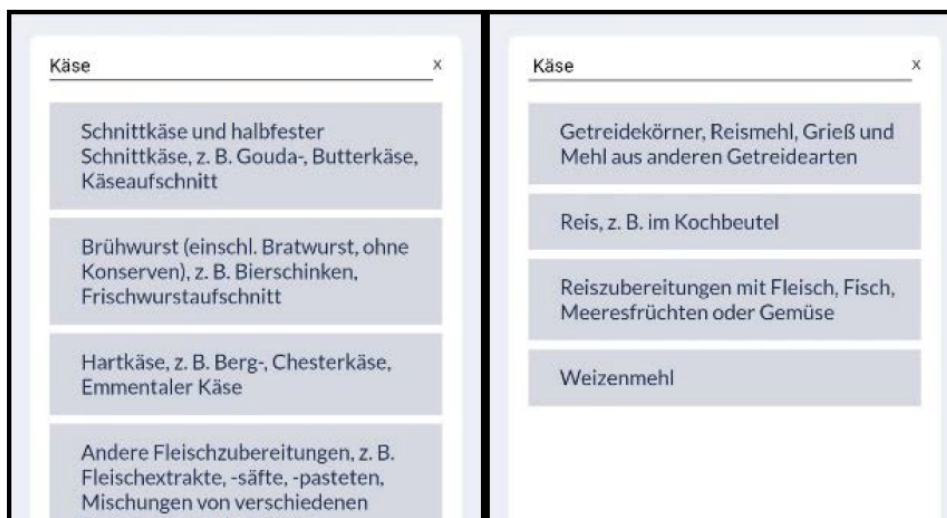
- o → Adding blank spaces at the end of a search term always leads to the display of the same four search results, completely independent of the actual term searched. The search algorithm should therefore ignore blank spaces if possible (or at least when blank spaces are not followed by additional text, i.e.: "Käse_" treated as "Käse")
- The suggestions"-functionality did not work.
  - o → The window that is supposed to list suggestions remains empty, even after several completed diary days full of expenditures entered into the app.
- The "filter"-functionality did not work:
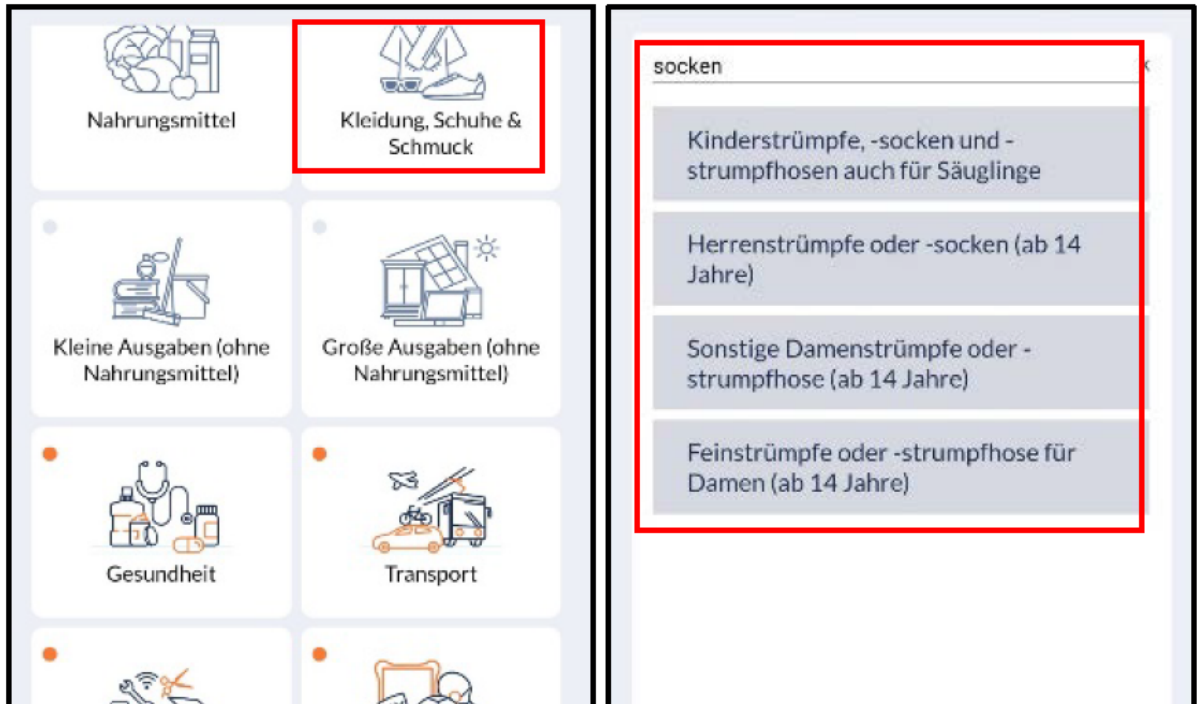
- o → Regardless of whether filters are selected or not, the search still displayed all expenditures: Deactivating the "clothing"-category (left screen – the category on the top left is not ticked) and searching for "socks" still lists search results clearly belonging to the "clothing"-category (all 4 search results shown).
- o We have tested it with different search terms and categories, no difference there (e.g.: banana and only filtering for food did not properly filter either).
- Long words are partly cut off:
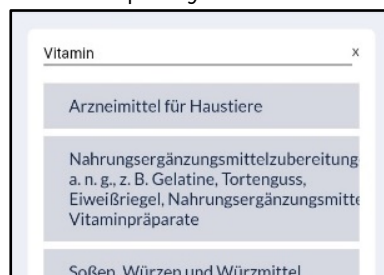
- o → Example: The German word „Nahrungsergänzungsmittelzubereitung" – it is getting cut off in-between, as the word is too long. In such cases, those responsible for textual input should be aware of a maximum number of characters per word. If a word exceeds that, breakpoints for a new paragraph should be defined.

Returns: Discounts, deposits, and other types

- If the discount is larger than the price of the product, the app gets stuck in a permanent "refreshing"-loop, and has to be restarted to function again:

- When simply returning deposit, the app "made up" a purchase of the amount and then subtracted it again. In case a person only returned their deposit and got money back, the app's calculation is wrong. The function that the amount of deposit is added first and then subtracted should be eliminated. Also in this case the three lower tiles do not make sense:

STATBEL
Belgium in figures

D STATIS
wissen.nutzen.

VUB VRIJE UNIVERSITEIT BRUSSEL

h

eurostat

Annex 3: WP4 – D4.2 – Report on the usability tests from Statbel

Report by Statbel – Hannelore Van der Beken, Patrick Lusyne

Background

We face a declining willingness of households to participate in the household budget survey. The response rate in the sample from the national register is alarmingly low - falling below 5%. So all hands on deck!

A survey such as the household budget survey is very time-consuming for respondents, especially as they have to keep detailed records of their income and expenditure over a period of time in the form of a diary. Since 2014, households have had the choice between a paper diary and a digital diary. The existing web application mainly eases the burden on the Citizens Survey Data Collection service in processing incoming data, but it does not reduce the workload for respondents. There are also other drawbacks to the existing tool. It is not straightforward to enter expenses on the go, potentially leading to underreporting of smaller expenses such as a take-home coffee or sweets from a vending machine. In addition, there is only one login per household. The expenditure of all household members must be recorded in the same diary. This may also lead to underreporting of certain expenses that someone does not want household members to know about.

It's time for something new! To reduce the workload for respondents and improve the quality of data collection, we are pinning our hopes on the MOTUS software platform designed by the TOR research group at the Free University of Brussels in the context of the time expenditure survey. Respondents will be able to enter their consumer spending data in a simple and convenient way using any device, either at home on their desktop PC or on the go on their mobile devices (tablet, smartphone). Innovative technologies such as receipt scanning are currently being developed and will be built into the platform. It is expected that the use of these new technologies, will provide households with an attractive and time-saving survey tool and make them more willing to participate in the household budget survey.

Intent

The development of the MOTUS software platform is still in progress. The current test is therefore a test of a product under development that has a number of limitations, such as incomplete shop and product lists and the absence of the household level.

This first test can be seen as a user acceptance test aimed at seeing whether the user can work with the software platform and whether the platform adequately facilitates the user, for whom it is intended.

Research flow

The test under the e-HBS project consists of three tasks.

1. Extensive testing of the cost diary over a number of days.
2. Completing a questionnaire on the home, energy and water expenses, insurance expenses, benefits in kind and purchase of durable goods (i.e. a shortened version of the household questionnaire normally administered by an interviewer).

3. Completing a questionnaire that gauges your experience with and opinion of the MOTUS application (installing the app, using the web application, login procedure, instructions, questioning, etc.).

The responses to the evaluation questionnaire will be used as a starting point for a focus group interview that will delve deeper into experiences with the MOTUS software platform.

The flow for the HBS diary CROESS Statbel
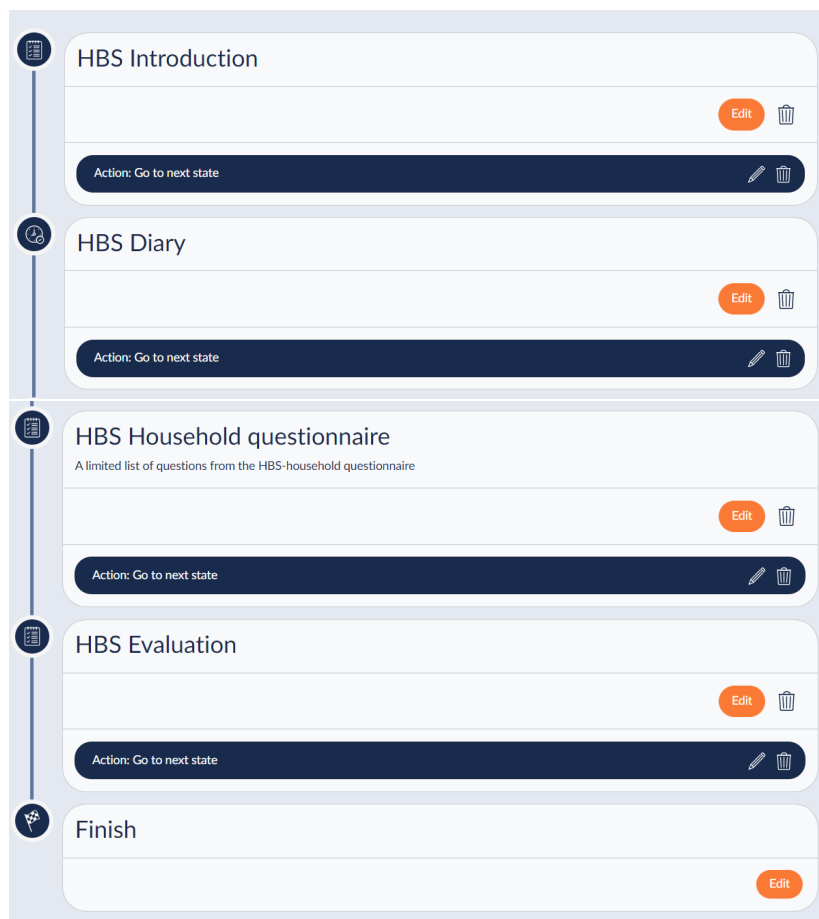
Communications

The invitation to participate in the test was sent by e-mail. Automated communication in the MOTUS platform was not used as it was outside the scope of this test.

Invitation by e-mail

> Dear "First_name",
>
> Thank you for participating in the test of the e-HBS project.
>
> We face a declining willingness of households to participate in the household budget survey. The response rate in the sample from the national register is alarmingly low - falling below 5%. So all hands on deck!
>
> A survey such as the household budget survey is very time-consuming for respondents, especially as they have to keep detailed records of their income and expenditure over a period of time in the form of a diary. Since 2014, households have had the choice between a paper diary and a digital diary. The existing web application mainly eases the burden

on the Citizens Survey Data Collection service in processing incoming data, but it does not reduce the workload for respondents. There are also other drawbacks to the existing tool. It is not straightforward to enter expenses on the go, potentially leading to underreporting of smaller expenses such as a take-home coffee or sweets from a vending machine. In addition, there is only one login per household. The expenditure of all household members must be recorded in the same diary. This may also lead to underreporting of certain expenses that someone does not want household members to know about.

It's time for something new! To reduce the workload for respondents and improve the quality of data collection, we are pinning our hopes on the MOTUS software platform designed by the TOR research group at the Free University of Brussels in the context of the time expenditure survey. Respondents will be able to enter their consumer spending data in a simple and convenient way using any device, either at home on their desktop PC or on the go on their mobile devices (tablet, smartphone). Innovative technologies such as receipt scanning are currently being developed and will be built into the platform. It is expected that the use of these new technologies, will provide households with an attractive and time-saving survey tool and make them more willing to participate in the household budget survey.

We are not at that point yet. Nevertheless, we would already like to appeal to you to test the tool. It is important to keep in mind here that it is not a finished product. The existing tool suffers from a number of limitations such as incomplete shop and product lists, the lack of the household level, ...

We like to hear afterwards what went well, but especially what could be improved. This is the only way we can further improve the research tool. If you should come across impeding bugs during testing, we ask you to point them out immediately by taking a screenshot and sending it to statbel.testpanel@economie.fgov.be.

How does it work?

The test under the e-HBS project consists of three tasks.

1. Extensive testing of the cost diary over several days. (! Give us a call if you finish testing the cost diary sooner, and we will activate the next phase of the test).
2. Completing a questionnaire on the home, energy and water expenses, insurance expenses, benefits in kind and purchase of durable goods (i.e. a shortened version of the household questionnaire normally administered by an interviewer).
3. Completion of a questionnaire probing your experience with and opinion of the MOTUS application (installing the app, using the web application, the login procedure, the instructions, the questioning, ...). The answers to this questionnaire will be used as a starting point for a focus group interview where we will go even deeper into your experiences with the new tool.

Ready to launch?
You can keep the cost diary and complete the questionnaires on the MOTUS website via your computer, laptop or tablet and via the MOTUS Discovery app on your smartphone or tablet.

Participate via the app - download the MOTUS Discovery application on your smartphone from the Apple App Store or from the Google Play Store, depending on your operating system

Participate via the website https://demo1-frontoffice.motusdemo.com/

Evaluation questionnaire

The evaluation questionnaire asks about the experience with and opinion on some technical aspects of the MOTUS application (installing the app, using the web application, the login procedure, the language setting, …) and this is followed by a series of questions on rather substantive aspects of the household budget survey test (the instructions, the question wording, …).

Number	Question & answer modality	Routing
	MOTUS app	
Q1.1	**Do you have the MOTUS app installed on your smartphone or tablet?**	
	Yes, via the Play Store (Google)	go to Q1.4
	Yes, via the App store (Apple)	go to Q1.4
	No, installing the MOTUS app failed	go to Q1.2
	No, only participated via the MOTUS web application	go to Q1.1_bis
Q1.1_bis	**Can you briefly say a little more about why you did not install the MOTUS app?**	
	Open field (short text) -- not required	go to Q2.4
Q1.2	**Why was installing the MOTUS app unsuccessful?**	
	Problems with internet connection	go to Q2.4
	Too little available storage space	go to Q2.4
	Outdated operating system	go to Q2.4
	MOTUS app not found	go to Q2.4
	Other reason	go to Q1.3
	No idea	go to Q2.4
Q1.3	**Can you briefly explain a little more why installing the MOTUS app failed?**	
	Open field (short text) -- not required	go to Q2.4
Q1.4	**Did you encounter any problems while installing the MOTUS app?**	
	Yes	go to Q1.5
	No	go to Q1.7
Q1.5	**What problems did you encounter while installing the MOTUS app?**	
	Problems with internet connection	go to Q1.7
	Too little available storage space	go to Q1.7
	Outdated operating system	go to Q1.7
	Other problems	go to Q1.6
Q1.6	**Can you briefly elaborate on the problems while installing the MOTUS app?**	
	Open field (short text) -- not required	go to Q1.7
Q1.7	**Did you experience any problems with the MOTUS app during the test (app crashed, stopped opening, was unresponsive, …)?**	
	Yes	go to Q1.8
	No	go to Q1.14
Q1.8	**How often did it happen that the MOTUS app crashed or did not work as expected?**	
	Continuing	go to Q1.9
	Regular	go to Q1.9

	Occasionally	go to Q1.9
	A single	go to Q1.9
Q1.9	**Can you briefly elaborate on the problems with the MOTUS app?**	
	Open field (short text) -- not required	go to Q1.10
Q1.10	**Have you reported the problems with the MOTUS app (e.g. by taking a screenshot and forwarding it to statbel.testpanel@economie.fgov.be)?**	
	Yes	go to Q1.11
	No	go to Q1.11
Q1.11	**Did the problems with the MOTUS app get resolved?**	
	Yes	go to Q1.14
	No	go to Q1.12
Q1.12	**Were the problems with the MOTUS app blocking further testing?**	
	Yes	go to Q1.13
	No	go to Q1.14
Q1.13	**Can you briefly elaborate on how you proceeded with the test?**	
	Open field (short text) -- not required	go to Q1.14
Q1.14	**How user-friendly do you find the MOTUS app?**	
	Extremely user-friendly	go to Q1.15
	Very user-friendly	go to Q1.15
	Reasonably user-friendly	go to Q1.15
	Not very user-friendly	go to Q1.15
	Not user-friendly at all	go to Q1.15
Q1.15	**How do you rate the MOTUS app in general?**	
	Very good	go to Q1.16
	Good	go to Q1.16
	Fair	go to Q1.16
	Bad	go to Q1.16
	Very bad	go to Q1.16
Q1.16	**In what areas could the MOTUS app be improved?**	
	Open field (long text) -- not required	go to Q2.1
	MOTUS website	
Q2.1	**Have you used the MOTUS web application?**	
	Yes	go to Q2.2
	No	go to Q3.1
Q2.2	**The MOTUS app and the MOTUS web application are aligned, making it possible to use both alternately. Have you taken advantage of this?**	
	Yes	go to Q2.3
	No	go to Q2.3
Q2.3	**Do you have a preference for either application?**	
	Preferred MOTUS app	go to Q2.3_bis
	Preferred MOTUS web application	go to Q2.3_tris
	No preference	go to Q2.4
Q2.3_bis	**Can you briefly say a little more about why you prefer the MOTUS app?**	
	Open field (short text) -- not required	go to Q2.4
Q2.3_tris	**Can you briefly say a little more about why you prefer the MOTUS web application?**	
	Open field (short text) -- not required	go to Q2.4
Q2.4	**Which web browser(s) did you use to reach the MOTUS web application?**	
	Google Chrome	go to Q2.5
	Internet Explorer	go to Q2.5
	Microsoft Edge	go to Q2.5
	Mozilla Firefox	go to Q2.5

	Opera	go to Q2.5
	Safari	go to Q2.5
	Other web browser	go to Q2.5
Q2.5	**Did you encounter problems with the MOTUS web application in a particular web browser?**	
	Yes	go to Q2.6
	No	go to Q2.7
Q2.6	**In which web browser did you have problems with the MOTUS web application?**	
	Open field (short text) -- not required	go to Q2.7
Q2.7	**How user-friendly do you find the MOTUS web application?**	
	Extremely user-friendly	go to Q2.8
	Very user-friendly	go to Q2.8
	Reasonably user-friendly	go to Q2.8
	Not very user-friendly	go to Q2.8
	Not user-friendly at all	go to Q2.8
Q2.8	**How do you rate the MOTUS web application in general?**	
	Very good	go to Q2.9
	Good	go to Q2.9
	Fair	go to Q2.9
	Bad	go to Q2.9
	Very bad	go to Q2.9
Q2.9	**In what areas can the MOTUS web application be improved?**	
	Open field (long text) -- not required	go to Q3.1
	MOTUS application	
Q3.1	**Did you experience difficulties logging in to the MOTUS application?**	
	Yes	go to Q3.2
	No	go to Q3.4
Q3.2	**How often did you experience difficulties logging in to the MOTUS application?**	
	Continuing	go to Q3.3
	Regular	go to Q3.3
	Occasionally	go to Q3.3
	A single	go to Q3.3
Q3.3	**What problems did you have logging in to the MOTUS application?**	
	Open field (short text) -- not required	go to Q3.4
Q3.4	**Did you experience difficulties moving from task to task in the MOTUS application (e.g. completing a questionnaire first, then completing a diary of daily expenses, then going back to a questionnaire...)?**	
	Yes	go to Q3.5
	No	go to Comm1
Q3.5	**How often did you experience difficulties moving from task to task in the MOTUS application?**	
	Continuing	go to Q3.6
	Regular	go to Q3.6
	Occasionally	go to Q3.6
	A single	go to Q3.6
Q3.6	**What problems did you have moving from task to task in the MOTUS application?**	
	Open field (short text) -- not required	go to Comm1
	MOTUS household budget survey	
Comm1	The household budget survey consisted of several tasks. We will go over these one by one.	go to Q4.1
Q4.1	**When opening the MOTUS cost diary, it was clear what was expected of you.**	

	Totally agree	go to Q4.3
	Agreed	go to Q4.3
	Disagree	go to Q4.2
	Totally disagree	go to Q4.2
Q4.2	**What ambiguities were there?**	
	Open field (short text) -- not required	go to Q4.4
Q4.3	**A short instructional video explaining step-by-step how to complete the cost diary is appropriate.**	
	Totally agree	go to Q4.4
	Agreed	go to Q4.4
	Disagree	go to Q4.4
	Totally disagree	go to Q4.4
Q4.4	**Overall, how do you rate the summary page in the cost diary?**	
	Very good	go to Q4.5
	Good	go to Q4.5
	Fair	go to Q4.5
	Bad	go to Q4.5
	Very bad	go to Q4.5
Q4.5	**What do you like about the overview page?**	
	Open field (short text) -- not required	go to Q4.6
Q4.6	**What don't you like about the overview page?**	
	Open field (short text) -- not required	go to Q4.7
Q4.7	**Do you think something is missing from the cost diary summary page?**	
	Yes	go to Q4.8
	No	go to Q4.9
Q4.8	**What do you think is missing from the overview page?**	
	Open field (short text) -- not required	go to Q4.9
Q4.9	**How do you rate the page structure, the arrangement of the different elements on the overview page?**	
	Very good	go to Q4.10
	Good	go to Q4.10
	Fair	go to Q4.10
	Bad	go to Q4.10
	Very bad	go to Q4.10
Q4.10	**In what areas could the structure of the overview page be improved?**	
	Open field (short text) -- not required	go to Comm2
Comm2	**The main component and purpose of the MOTUS application is to enter all expenses. That process will now be gone over step by step.**	go to Q4.11
Q4.11	**It was immediately clear on the overview page how to enter a new issue.**	
	Totally agree	go to Q4.12
	Agreed	go to Q4.12
	Disagree	go to Q4.12
	Totally disagree	go to Q4.12
Q4.12	**It was immediately clear on the overview page that monthly fixed charges had to be entered through a separate procedure.**	
	Totally agree	go to Q4.13
	Agreed	go to Q4.13
	Disagree	go to Q4.13
	Totally disagree	go to Q4.13
Q4.13	**It was immediately clear on the summary page how to enter that you had spent nothing on a particular day.**	
	Totally agree	go to Comm3
	Agreed	go to Comm3
	Disagree	go to Comm3

	Totally disagree	go to Comm3
Comm3	**On a new release, you first got an input screen where you had to enter the date, country and place of purchase.**	
Q4.14	**Did you experience difficulties entering the date?**	
	Yes	go to Q4.15
	No	go to Q4.17
Q4.15	**How often did you experience difficulties entering the date?**	
	Continuing	go to Q4.16
	Regular	go to Q4.16
	Occasionally	go to Q4.16
	A single	go to Q4.16
Q4.16	**What problems did you have entering the date?**	
	Open field (short text) -- not required	go to Q4.17
Q4.17	**Did you experience difficulties entering the country of purchase?**	
	Yes	go to Q4.18
	No	go to Q4.20
Q4.18	**How often did you experience difficulties entering the country of purchase?**	
	Continuing	go to Q4.19
	Regular	go to Q4.19
	Occasionally	go to Q4.19
	A single	go to Q4.19
Q4.19	**What problems did you have entering the country of purchase?**	
	Open field (short text) -- not required	go to Q4.20
Q4.20	**Did you experience difficulties entering the place of purchase?**	
	Yes	go to Q4.21
	No	go to Comm4
Q4.21	**How often did you experience difficulties entering the place of purchase?**	
	Continuing	go to Q4.22
	Regular	go to Q4.22
	Occasionally	go to Q4.22
	A single	go to Q4.22
Q4.22	**What problems did you have entering the place of purchase?**	
	Open field (short text) -- not required	go to Comm4
Comm4	**You were then presented with an input screen where you had to enter detailed information about the products/services purchased (description, category, number, price, quantity, unit and any discounts).**	go to Q4.23
Q4.23	**Did you experience difficulties entering the category of products/services purchased?**	
	Yes	go to Q4.24
	No	go to Q4.26
Q4.24	**How often did you experience difficulties entering the category of products/services purchased?**	
	Continuing	go to Q4.25
	Regular	go to Q4.25
	Occasionally	go to Q4.25
	A single	go to Q4.25
Q4.25	**What problems did you have entering the category of products/services purchased?**	
	Open field (short text) -- not required	go to Q4.26
Q4.26	**Did you experience difficulties entering the number?**	
	Yes	go to Q4.27
	No	go to Q4.29
Q4.27	**How often did you experience difficulties entering the number?**	

	Continuing	go to Q4.28
	Regular	go to Q4.28
	Occasionally	go to Q4.28
	A single	go to Q4.28
Q4.28	**What problems did you have entering the number?**	
	Open field (short text) -- not required	go to Q4.29
Q4.29	**Did you experience difficulties entering the price?**	
	Yes	go to Q4.30
	No	go to Q4.32
Q4.30	**How often did you experience difficulties entering the price?**	
	Continuing	go to Q4.31
	Regular	go to Q4.31
	Occasionally	go to Q4.31
	A single	go to Q4.31
Q4.31	**What problems did you have entering the price?**	
	Open field (short text) -- not required	go to Q4.32
Q4.32	**Did you experience difficulties entering the quantity?**	
	Yes	go to Q4.33
	No	go to Q4.35
Q4.33	**How often did you experience difficulties entering the quantity?**	
	Continuing	go to Q4.34
	Regular	go to Q4.34
	Occasionally	go to Q4.34
	A single	go to Q4.34
Q4.34	**What problems did you have entering the quantity?**	
	Open field (short text) -- not required	go to Q4.35
Q4.35	**Did you experience difficulties in entering the unit?**	
	Yes	go to Q4.36
	No	go to Q4.38
Q4.36	**How often did you experience difficulties entering the unit?**	
	Continuing	go to Q4.37
	Regular	go to Q4.37
	Occasionally	go to Q4.37
	A single	go to Q4.37
Q4.37	**What problems did you have entering the unit?**	
	Open field (short text) -- not required	go to Q4.38
Q4.38	**Did you experience difficulties entering discounts?**	
	Yes	go to Q4.39
	No	go to Q4.41
Q4.39	**How often did you experience difficulties entering discounts?**	
	Continuing	go to Q4.40
	Regular	go to Q4.40
	Occasionally	go to Q4.40
	A single	go to Q4.40
Q4.40	**What problems did you have entering discounts?**	
	Open field (short text) -- not required	go to Q4.41
Q4.41	**The individual steps (search product/service first and then enter number, price, quantity, unit and any discounts) are logically arranged.**	
	Totally agree	go to Q4.42
	Agreed	go to Q4.42
	Disagree	go to Q4.42
	Totally disagree	go to Q4.42
Q4.42	**It was immediately clear how to enter several purchases during the same shop visit.**	

	Totally agree	go to Q4.43
	Agreed	go to Q4.43
	Disagree	go to Q4.43
	Totally disagree	go to Q4.43
Q4.43	**It is easy to enter monthly fixed costs in the cost diary.**	
	Totally agree	go to Q4.44
	Agreed	go to Q4.44
	Disagree	go to Q4.44
	Totally disagree	go to Q4.44
Q4.44	**Did you experience difficulties entering monthly fixed costs?**	
	Yes	go to Q4.45
	No	go to Q4.47
Q4.45	**How often did you experience difficulties entering monthly fixed costs?**	
	Continuing	go to Q4.46
	Regular	go to Q4.46
	Occasionally	go to Q4.46
	A single	go to Q4.46
Q4.46	**What problems did you have entering monthly fixed costs?**	
	Open field (short text) -- not required	go to Q4.47
Q4.47	**What was your overall experience with the cost diary in the MOTUS application?**	
	Very good	go to Q4.48
	Good	go to Q4.48
	Fair	go to Q4.48
	Bad	go to Q4.48
	Very bad	go to Q4.48
Q4.48	**What did you like?**	
	Open field (short text) -- not required	go to Q4.49
Q4.49	**What did you think was bad?**	
	Open field (short text) -- not required	go to Q4.50
Q4.50	**Do you have a preference for the cost diary on paper or the cost diary in the MOTUS application?**	
	Preferred paper-based cost diary	go to Q4.51
	Preference for the cost diary in the MOTUS application	go to Q4.52
	No preference	go to Comm4
Q4.51	**Can you briefly elaborate on why you prefer the cost diary on paper?**	
	Open field (max 300 characters) -- not required	go to Comm4
Q4.52	**Can you briefly elaborate on why you prefer the cost diary in the MOTUS application?**	
	Open field (max 300 characters) -- not required	go to Comm4
Comm4	**The final task in the household budget survey involved answering a questionnaire that covered your home, your energy and water expenses, your insurance expenses and some major purchases you have made in the past few months (i.e. a shortened version of the household questionnaire normally administered by an interviewer).**	go to Q4.53
Q4.53	**Answering the questions about your living situation was easy.**	
	Totally agree	go to Q4.55
	Agreed	go to Q4.55
	Disagree	go to Q4.54
	Totally disagree	go to Q4.54
Q4.54	**What difficulties were there?**	
	Open field (max 300 characters) -- not required	go to Q4.55
Q4.55	**Answering the questions on periodic expenses related to your home (energy and water expenses) was easy.**	
	Totally agree	go to Q4.57

	Agreed	go to Q4.57
	Disagree	go to Q4.56
	Totally disagree	go to Q4.56
Q4.56	**What difficulties were there?**	
	Open field (max 300 characters) -- not required	go to Q4.57
Q4.57	**Periodic expenses related to your home could also be entered under fixed monthly expenses in the cost diary. Do you have a preference for including such expenses in the cost diary or do you prefer the retrospective questionnaire?**	
	Preferred all-in cost diary	go to Q4.58
	Preferred questionnaire afterwards	go to Q4.58
	No preference	go to Q4.58
Q4.58	**Answering the questions on the goods owned by your household (vehicle and mobile phone) was easy.**	
	Totally agree	go to Q4.60
	Agreed	go to Q4.60
	Disagree	go to Q4.59
	Totally disagree	go to Q4.59
Q4.59	**What difficulties were there?**	
	Open field (max 300 characters) -- not required	go to Q4.60
Q4.60	**Expenses related to goods owned by your household could also be entered under fixed monthly expenses in the cost diary. Do you have a preference for including such expenses in the cost diary or do you prefer the retrospective questionnaire?**	
	Preferred all-in cost diary	go to Q4.61
	Preferred questionnaire afterwards	go to Q4.61
	No preference	go to Q4.61
Q4.61	**Answering the questions on purchasing durable goods (major and minor household appliances) was easy.**	
	Totally agree	go to Q4.63
	Agreed	go to Q4.63
	Disagree	go to Q4.62
	Totally disagree	go to Q4.62
Q4.62	**What difficulties were there?**	
	Open field (max 300 characters) -- not required	go to Q4.63
Q4.63	**Does it seem appropriate to you to include the purchase of durable goods in the cost diary in the MOTUS application (by analogy with monthly fixed costs)?**	
	Yes	go to Q4.64
	No	go to Q4.64
Q4.64	**Answering the questions on the insurance policies taken out by your household was easy.**	
	Totally agree	go to Q4.66
	Agreed	go to Q4.66
	Disagree	go to Q4.65
	Totally disagree	go to Q4.65
Q4.65	**What difficulties were there?**	
	Open field (max 300 characters) -- not required	go to Q4.66
Q4.66	**Expenses related to insurance policies taken out by your household could also be entered under fixed monthly expenses in the cost diary. Do you have a preference for including such expenses in the cost diary or do you prefer the retrospective questionnaire?**	
	Preferred all-in cost diary	go to Q4.67
	Preferred questionnaire afterwards	go to Q4.67
	No preference	go to Q4.67

Q4.67	**Answering the questions on in-kind benefits received by your household was straightforward.**	
	Totally agree	go to Q4.69
	Agreed	go to Q4.69
	Disagree	go to Q4.68
	Totally disagree	go to Q4.68
Q4.68	**What difficulties were there?**	
	Open field (max 300 characters) -- not required	go to Q4.69
Q4.69	**Does it seem appropriate to you to include benefits in kind in the cost diary in the MOTUS application (by analogy with monthly fixed costs)?**	
	Yes	go to Q4.70
	No	go to Q4.70
Q4.70	**Any comments, questions or suggestions?**	
	Open field (max 300 characters) -- not required	end

19.1 Test panel

Six Dutch-speaking Statbel employees were invited to participate in the test.

Three HBS experts testing based on their experience and three staff without specific knowledge about HBS but with experience in data collection. The latter were provided with a paper cost diary before the start of the test in order to compare the different modes of data collection afterwards.

Used devices

The information regarding the devices used was manually extracted from the MOTUS back office for the 6 respondents.

Data on the devices used per survey phase is not available. Consequently, it cannot be verified whether there is a difference in preference/use between the MOTUS web application (accessed via computer, laptop or tablet) and the MOTUS app (on smartphone or tablet) in terms of completing the questionnaires vs keeping the cost diary.
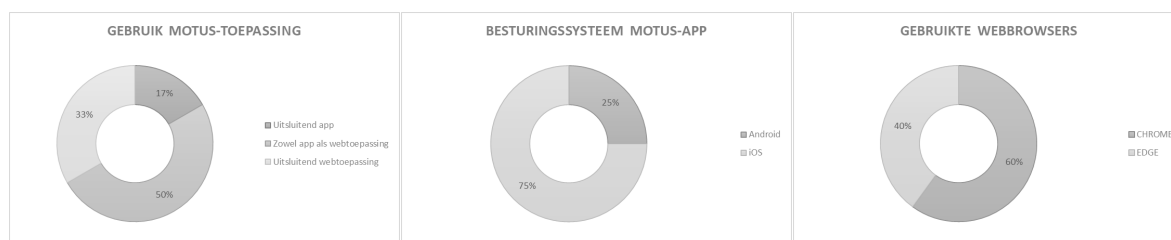
Only one respondent participated exclusively through the app. Half of the respondents (N=3) used both the app and the web application. The remaining two respondents participated exclusively through the website.

Three app users have an iPhone (with operating system iOS 15 or iOS 16). The fourth app user installed the MOTUS app on an Android smartphone (with operating system Android 8).

Our respondents appear to be very browser loyal. No one used multiple web browsers to access the MOTUS web application. The most frequently used browser is Chrome, not surprising given that it is currently the most widely used web application in the world. One respondent mentioned that it was not possible to reach the MOTUS web application through Safari.

With the exception of one respondent who occasionally had problems logging in via the app, no other major problems were reported with the application in general.

The MOTUS application in the e-HBS project

Summary page in the cost diary

For one respondent, when opening the MOTUS cost diary, it was not at all clear what was expected.

> "*Filling in what where and most importantly, how does it work?* " (Respondent fb37c5b7-7c35-4d21-b319-6a2045b82e1c)

For the remaining respondents, there were no ambiguities in this respect, but all felt that a short instructional video explaining step-by-step how to fill in the cost diary would be appropriate.

The summary page in the cost diary is generally evaluated rather positively. One respondent explicitly mentions the listing of the various purchases as a plus. Another respondent felt that there might be too much on the page, making it not always clear what needs to be filled in. This is echoed by another respondent who feels that there is too little info to pull off.

STATBEL — Belgium in figures

DESTATIS — wissen.nutzen.

VUB VRIJE UNIVERSITEIT BRUSSEL

h

eurostat
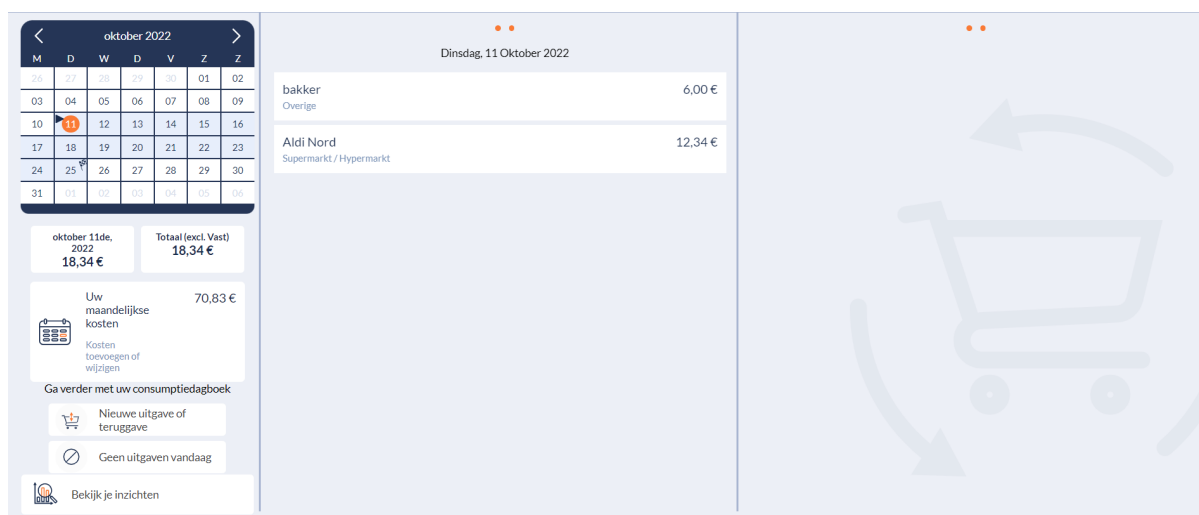
≡ CONSUMPTIEDAGBOEK 🛒

okt. 2022

M	D	W	D	V	Z	Z
26	27	28	29	30	01	02
03	04	05	06	07	08	09
10	11	12	13	14	15	16
17	18	19	20	21	22	23
24	25	26	27	28	29	30

11 oktober 2022
18,34 €

Totaal (excl. Vast)
18,34 €

Uw maandelijkse kosten — 70,83 €
Kosten toevoegen of wijzigen

Uitgaven : dinsdag 11 oktober

bakker — 6,00 €
Speciaalzaak

Aldi Nord — 12,34 €
Supermarkt / Hypermarkt

Inzichten

Belangrijkste uitgavencategorieën

■ Vaste kosten (66%)
■ Levensmiddelen (23%)
■ Cultuur, sport & vrije tijd (11%)

53,27 €

Statistieken tonen

Herladen

+
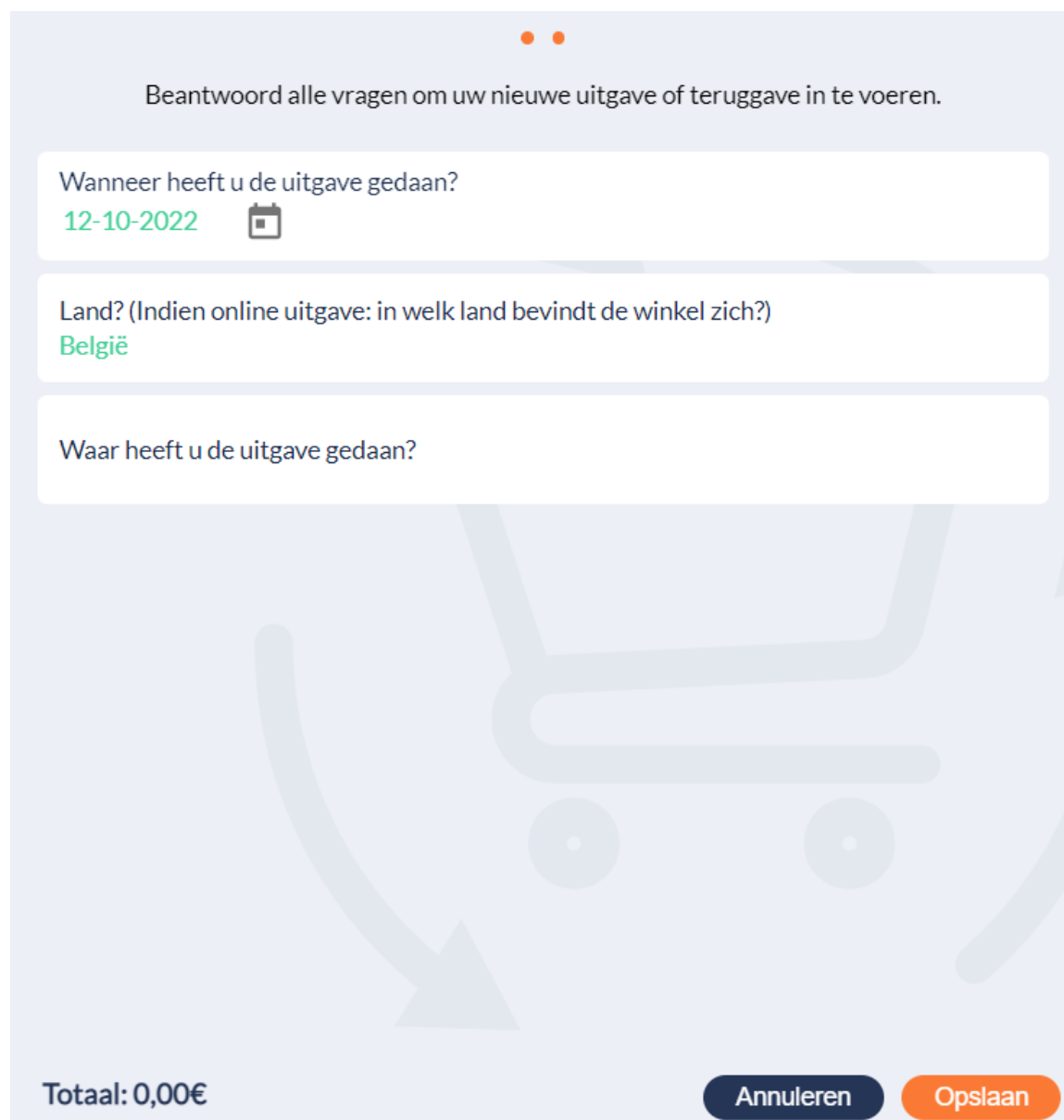
The page structure, i.e. the arrangement of the various elements on the overview page, is rated as good by the respondents. Nevertheless, two respondents indicated that it was not immediately clear on the overview page how to enter a new issue. For three respondents, it was not immediately clear on the summary page that monthly fixed charges had to be entered through a separate procedure. And for one respondent, it was also not immediately clear on the summary page how to enter that one had not spent anything on a particular day.

This points to the need for clear instructions. By analogy with the MOTUS application for the time spent survey, an "assistant" should also be available here to guide the respondent through the cost diary with notifications explaining the steps the respondent needs to take to record a purchase.

Entering expenses in the cost diary

The main component and purpose of the MOTUS application is to enter all editions. When entering a new issue, one first gets an entry screen where the date, country and place of purchase had to be entered - see Figure 85.

One respondent indicated that she regularly encountered difficulties in entering the date.

> *"Had to change the date of certain expenses a few times, either it was not remembered by the application, or it was just me and I had not properly understood how it worked"* (Respondent fb37c5b7-7c35-4d21-b319-6a2045b82e1c)

Another respondent pointed out that everything hinges on the date. If the exact date of the purchase is no longer known, the purchase cannot be registered.

No respondents experienced difficulties entering the country of purchase. Entering the place of purchase, on the other hand, did present a lot of difficulties.

> *"The place of purchase was not listed - I did not know how to name the shop according to the descriptions given."* (Respondent fb37c5b7-7c35-4d21-b319-6a2045b82e1c)

> *"The search engine did not find the place. It didn't find bakery or butcher shop, for example."* (Respondent 58195a7a-ce4a-440c-92b4-1949730bc6bd)

> *"Wrong language in the lookup list. "* (Respondent e875cd7e-e9e2-468c-b22d-749ebe8e1abe)
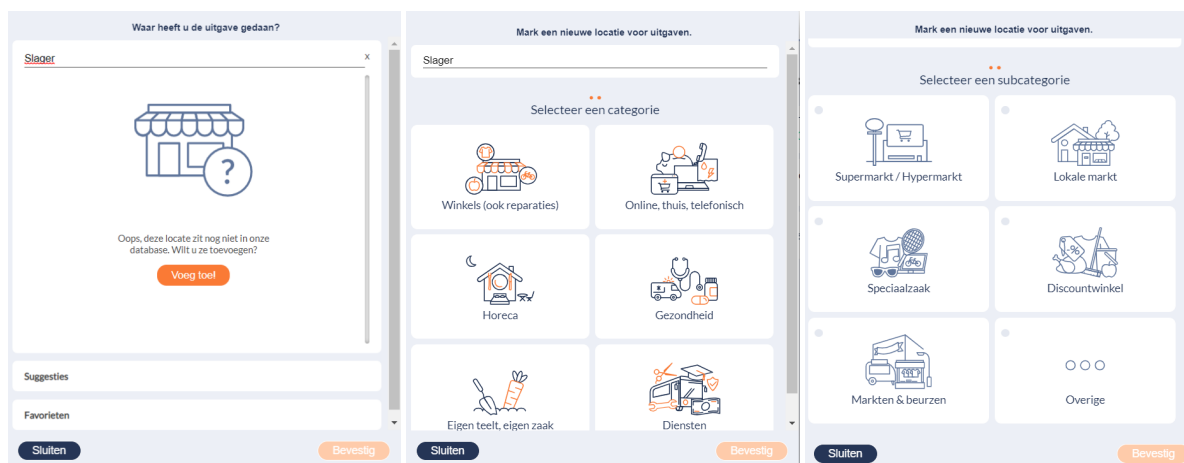
> *"Courant shops should add themselves under 'Other' category. '* (Respondent e875cd7e-e9e2-468c-b22d-749ebe8e1abe)

The reported problems were all due to an incomplete database that was also partly in German - see Figure 86. Respondents had been informed at the start of the test that the tool suffered from some limitations such as an incomplete shopping list, but we underestimated the problems this caused. Because common locations such as bakery, butcher, petrol station, … were not known, entering purchases became a time-consuming activity. Each unknown location had to be added to the database by the respondent himself by assigning a category and a subcategory - see Figure 87. Naturally, this raised a lot of questions among the respondents: Is the database updated only for themselves or for all users? Are added locations checked and corrected, if necessary, by the back office? …

A subsequent input screen required entering detailed information about the products/services purchased (description, category, number, price, quantity, unit and any discounts) - see figures 88 and 90.

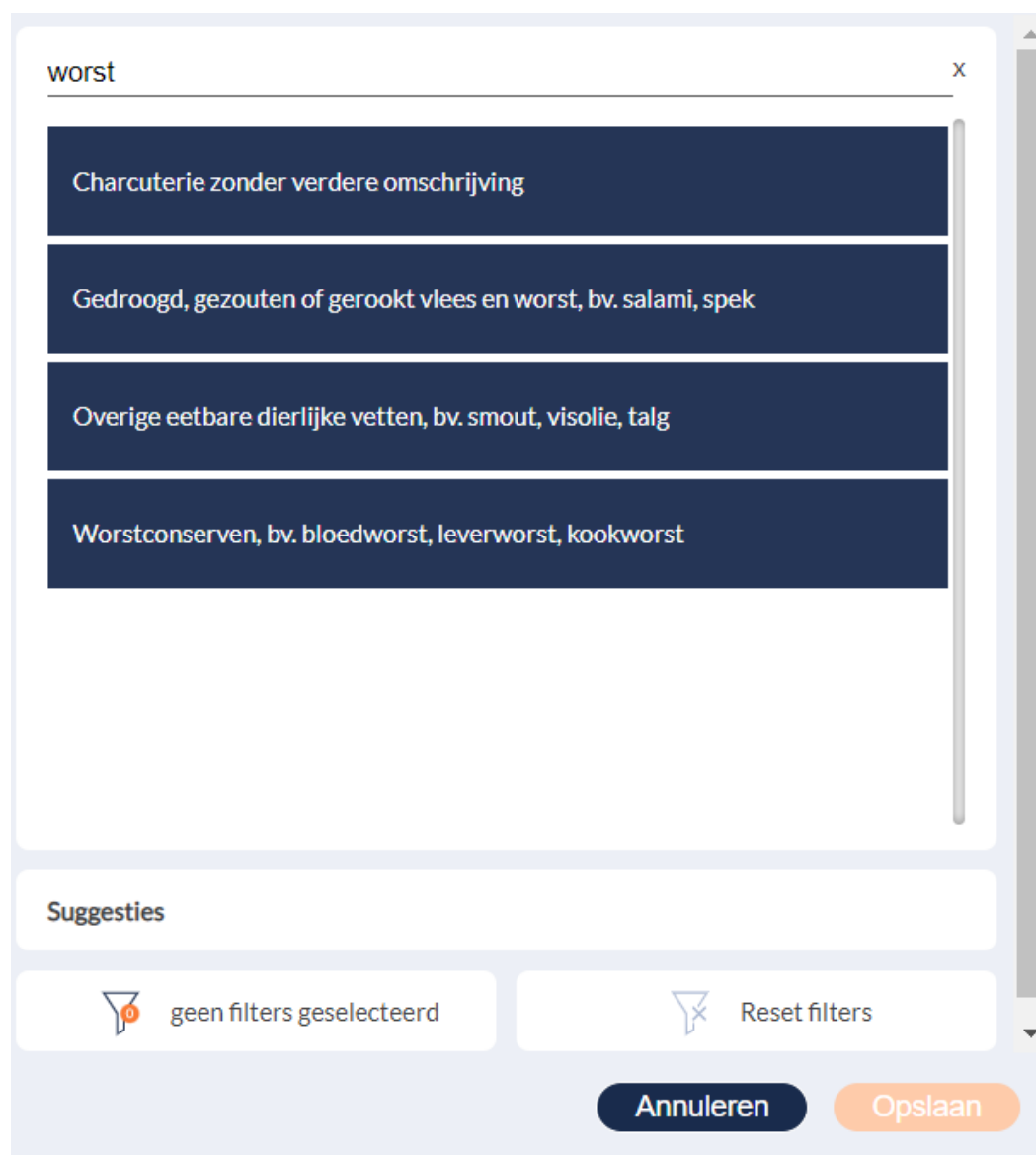A lot of difficulties were encountered in entering the category of products/services purchased.

> *"The categories provided do not always match the descriptions entered or are not always clear."* (Respondent fb37c5b7-7c35-4d21-b319-6a2045b82e1c)

> *"When you start typing the name of the product, suggestions appeared after only 2 letters. If you clicked on it, the system did not take the suggestion. For the description, the 2 letters remained. Then had to return to write the name of the product in full."* (Respondent 58195a7a-ce4a-440c-92b4-1949730bc6bd)

> *"I didn't find the right category."* (Respondent e875cd7e-e9e2-468c-b22d-749ebe8e1abe)

> *"It was not clear which category to choose. "* (Respondent e875cd7e-e9e2-468c-b22d-749ebe8e1abe)

Again, the reported problems are largely due to an incomplete database - see Figure 89. An obvious category such as 'fresh sausage' is missing, while a category such as 'other edible animal fats, e.g. smouth, fish oil, tallow' seems rather far-fetched when describing 'sausage' .

An additional problem was that, unlike the location database, the product list could not be completed by the respondent. To proceed, an incorrect category had to be indicated. This caused annoyance among respondents.

Beantwoord alle vragen om uw nieuwe uitgave of teruggave in te voeren.

Naam van de uitgave
worst

Soort uitgave
Gedroogd, gezouten of gerookt vlees en worst, bv. salami, spek

Hoeveelheid ___1___   Prijs ___8___ €

Aanvullende informatie (optioneel)

Bestaat uw uitgave uit meerdere artikelen?

○ Nee, het is 1 item

✓ Ja, ___4___ items

Meeteenheid van item(s) ___150___   Eenheid
g

Korting in euro? ___0___ €

Is dit een retourartikel? ○

Dit item uitsluiten van het huishoudbudget? Anderen zullen dit artikel niet zien ○

Totaal: 8,00€   Annuleren   Opslaan

Entering quantity, price, number, unit and any discounts presented fewer problems. Nevertheless, a number of justified comments were made about this too. Entering requires a relatively large amount of information per product - although not all fields are mandatory. One respondent commented that without a detailed receipt it is very difficult to display all the requested information.

> "Without detailed receipt there is no info on unit price, price/kg, number of pieces - you do know the price of 1 pack with x number of pieces. " (Respondent fb37c5b7-7c35-4d21-b319-6a2045b82e1c)

Due to the lack of clear instructions on what should be entered where, one of the HBS experts fears that people will write down certain information incorrectly, e.g. entering the number of litres of petrol filled in the quantity question and the full amount paid in the price question, which will lead to wrong results, i.e. an unintended multiplication of the effective price.

Two respondents mentioned difficulties in entering the price. Large amounts cannot be entered. A comma automatically appears as soon as an amount greater than € 999 is entered. It is not clear whether this is a deliberate choice or an unintentional error.

One respondent mentioned a problem in entering the unit.

> "When an item weighs 1.2 kg, you cannot put a comma. You have to solve this by taking 1200 grams the." (Respondent 58195a7a-ce4a-440c-92b4-1949730bc6bd)

Two respondents felt that the individual steps (searching for product/service first and then entering number, price, quantity, unit and any discounts) were not arranged logically. One respondent did not notice the option to enter discounts as a result.

> "Erm… discounts? I haven't come across that possibility. Maybe overlooked? " (Respondent fb37c5b7-7c35-4d21-b319-6a2045b82e1c)

For all but one respondent, it was also not immediately clear how to enter different purchases during the same shopping trip. This is yet another signal that clear instructions are indispensable to create a functioning application.

The possibility of not including a purchased item in the household budget - preventing housemates from seeing the purchase - could not be tested due to the lack of the household level.

19.1.1   Entering fixed costs in the cost diary

In the cost diary of the MOTUS application, it is also possible to enter fixed costs - see Figure 91 for a listing of the various fixed costs provisionally included.

Overzicht vaste kosten

● ●

Vul de vaste kosten in die op u van toepassing zijn.

Onderhoudsgelden Alimentatie, schooltoezicht,...		0/5
Huur en lasten Huur, gemeenschappelijke kosten, diensten...		0/17
Financiële diensten Bankkosten, persoonlijke leningen, hypotheek...		0/5
Nutsvoorzieningen Gas, elektriciteit, water, mazout...		0/12
Afval en riolering Kosten en taksen voor afvalverwerking en riolering		0/16
Mobiliteit Voertuigen, leasing, openbaar vervoer...		0/3
Telecommunicatie en media Abonnementen internet, telefonie, tv, kranten...		0/29
Cultuur en vrije tijd Lidmaatschap sportclub, bioscoop,...		1/4
Opleiding Opleidingsgelden, bijles...		0/1
Verzekeringen Aansprakelijkheid, brandverzekering, pechbijstand...		0/32

Maandelijks totaal: 70,83 €

Bevestigen

All respondents except one indicated that it was easy to enter monthly fixed costs in the cost diary. Two respondents nevertheless experienced some difficulties in entering monthly fixed costs. These difficulties are due to the fact that it is a hugely time-intensive task - both clicking on the various items to see the subcategories and looking it up to answer effectively.

> *"Difficulties I wouldn't call it, but a lot of searching if this is for an entire (large) household"* (Respondent fb37c5b7-7c35-4d21-b319-6a2045b82e1c)

Moreover, there were some annoyances because certain fixed cost items were missing and the frequency of payment could not be reflected correctly - quarterly, quarterly and six-monthly are missing from the response options.

Overall experience with the cost diary

Opinions are divided when asked about the overall experience with the cost diary in the MOTUS application. Two respondents with no previous HBS experience give a good rating, two HBS experts rate their experience as fair and the remaining two respondents are negative about the cost diary.

> *"The tool is structurally structured, but it is annoying that the search list does not find some products and you then have to rename them."* (Respondent 58195a7a-ce4a-440c-92b4-1949730bc6bd)

> *"I didn't find it easy, didn't go naturally, didn't work smoothly. It took me a while to get it right. It was time-consuming. I only worked through the PC (entering expenses at the end of the day/week). I did not use the app, not even while shopping. I can imagine that a visit to the shop can take longer than that. Not recommended for families with children and a busy schedule. "* (Respondent fb37c5b7-7c35-4d21-b319-6a2045b82e1c)

> *"Time-consuming. Lots of scrolling."* (Respondent e875cd7e-e9e2-468c-b22d-749ebe8e1abe)

> *"Unclear and very cumbersome."* (Respondent df91ac43-b434-4a67-8c86-db5c68481d31)

When respondents were asked about their preference - either the paper-based cost diary or the cost diary in the MOTUS application - only one respondent preferred the MOTUS application. Three respondents prefer the paper-based cost diary. The remaining two have no preference.

> *"Easy and quick to enter your daily expenses than on paper."* (Respondent 58195a7a-ce4a-440c-92b4-1949730bc6bd - preferred cost diary in MOTUS application)

> *"I have not worked on paper but am definitely not a fan of the application. On paper, it might be easier to define yourself (??) in the application you depend on the choices given, before you have found what you are looking for ..."* (Respondent fb37c5b7-7c35-4d21-b319-6a2045b82e1c - preference for cost diary on paper)

> *"Faster completed, everything on 1 line ... simpler"* (Respondent e875cd7e-e9e2-468c-b22d-749ebe8e1abe - preference for cost diary on paper)

Based on the feedback received, we have the impression that the resistance to the cost diary in the MOTUS application is mainly down to the incomplete and sometimes

untranslated search lists. We need to invest in an accurate and detailed database that does not lose sight of the day-to-day language.

In addition, the sequence of the different questions is not considered very intuitive and consistent. Clear instructions could address this, but perhaps the design also needs tinkering.

Respondents also want to know whether in the future it will be possible to adjust the currency in order to enter expenses during a trip abroad more easily. Will the MOTUS application be able to convert the foreign valuata so that these expenses can be correctly reflected in the total expenditure pattern?

As for innovative technologies, the question is raised whether voice input will be possible in addition to receipt scanning. The bar is set fairly high among frequent app users. Nowadays, there are many apps to choose from to help you manage your budget - see, for example, https://financer.com/nl/blog/huishoudboekje-app/ for a list of the best household bookkeeping apps. Based on experiences with such apps, expectations are created.

Household questionnaire

The last task in the test consisted of answering a questionnaire related to the home, energy and water expenses, insurance expenses and some major purchases one had made in the past few months. This was an abbreviated version of the household questionnaire normally administered by an interviewer.

Answering the questions on living situation was rated as easy by all respondents.

For all respondents, answering the questions on periodic home-related expenses (energy and water expenses) - see Figure 92 - was straightforward. Periodic home-related expenses could also be entered under the fixed monthly expenses in the cost diary - see Figure 93. Three respondents preferred to include such expenses in the cost diary. One respondent prefers the retrospective questionnaire. The remaining two have no preference for any particular mode.

Nu volgen enkele vragen over uw jaarlijkse water- en energieverbruik en enkele andere terugkerende uitgaven in verband met uw woning.

VRAAG 5

• •

Hoeveel betaalt u jaarlijks voor uw waterverbruik, voor de gemeentelijke bijdrage voor de afvoer van afvalwater en voor de bovengemeentelijke bijdrage voor de zuivering van afvalwater?

Ik ken de afzonderlijke bedragen voor waterverbruik, afvoer en zuivering van afvalwater

Ik ken enkel het gezamenlijk bedrag voor waterverbruik, afvoer en zuivering van afvalwater

Ik weet het niet, maar weet het wel voor verbruik van water, elektriciteit en gas samen

Ik weet het niet, maar weet het wel voor verbruik van water en elektriciteit samen

Ik weet het niet, maar weet het wel voor verbruik van water en gas samen

Ik weet het niet, want het bedrag is inbegrepen in de huurprijs

VRAAG 6

• •

Jaarlijks bedrag voor uw waterverbruik, voor de gemeentelijke bijdrage voor de afvoer van afvalwater en voor de bovengemeentelijke bijdrage voor de zuivering van afvalwater samen:

_____

• •

Nutsvoorzieningen

Waterrekening / distributie water / factuur voor hoofdverblijf

Waterrekening / distributie water / factuur voor tweede woning

Lasten huurder appartement

Lasten eigenaar appartement

Gezamenlijke diensten / kosten appartement: collectieve verwarming

Energie factuur hoofdverblijfplaats

Energie factuur tweede woning

Maandelijks totaal: 0,00 €   Terug   Bevestigen

Nutsvoorzieningen: Waterrekening / distributie water / factuur voor hoofdverblijf

Naam
Antwoord hier

Prijs in euro   0   €   Frequentie
                                    maandelijks

Professionele uitgave? ○   Betaald met creditcard? ○

Dit item uitsluiten van het huishoudbudget? Anderen zullen dit artikel niet zien ○

Toevoegen

Maandelijks totaal: 0,00 €

For all but one respondent, answering the questions on household owned goods (vehicle and mobile phone) was straightforward.

> "The composition of the 'packages' do not always match my situation. Either expand the choices (even more?) or be able to enter the composition of one's own package (by supplier/company) myself." (Respondent fb37c5b7-7c35-4d21-b319-6a2045b82e1c)

Expenditure related to goods owned by the household could also be entered under fixed monthly costs in the cost diary. Three respondents preferred the inclusion of such expenses

in the cost diary. One respondent prefers the retrospective questionnaire. The remaining two have no preference.

For all but one respondent, answering the questions on purchasing durable goods (large and small household appliances) was straightforward.

> "People are not going to remember the date of purchase of all these devices I'm afraid. " (Respondent df91ac43-b434-4a67-8c86-db5c68481d31)

Four respondents believe it is appropriate to include the purchase of durable goods in the cost diary in the MOTUS application (by analogy with monthly fixed costs).

For all but one respondent, answering the questions on the insurance policies taken out by the household was straightforward.

> "A lot of searching for the families …. " (Respondent df91ac43-b434-4a67-8c86-db5c68481d31)
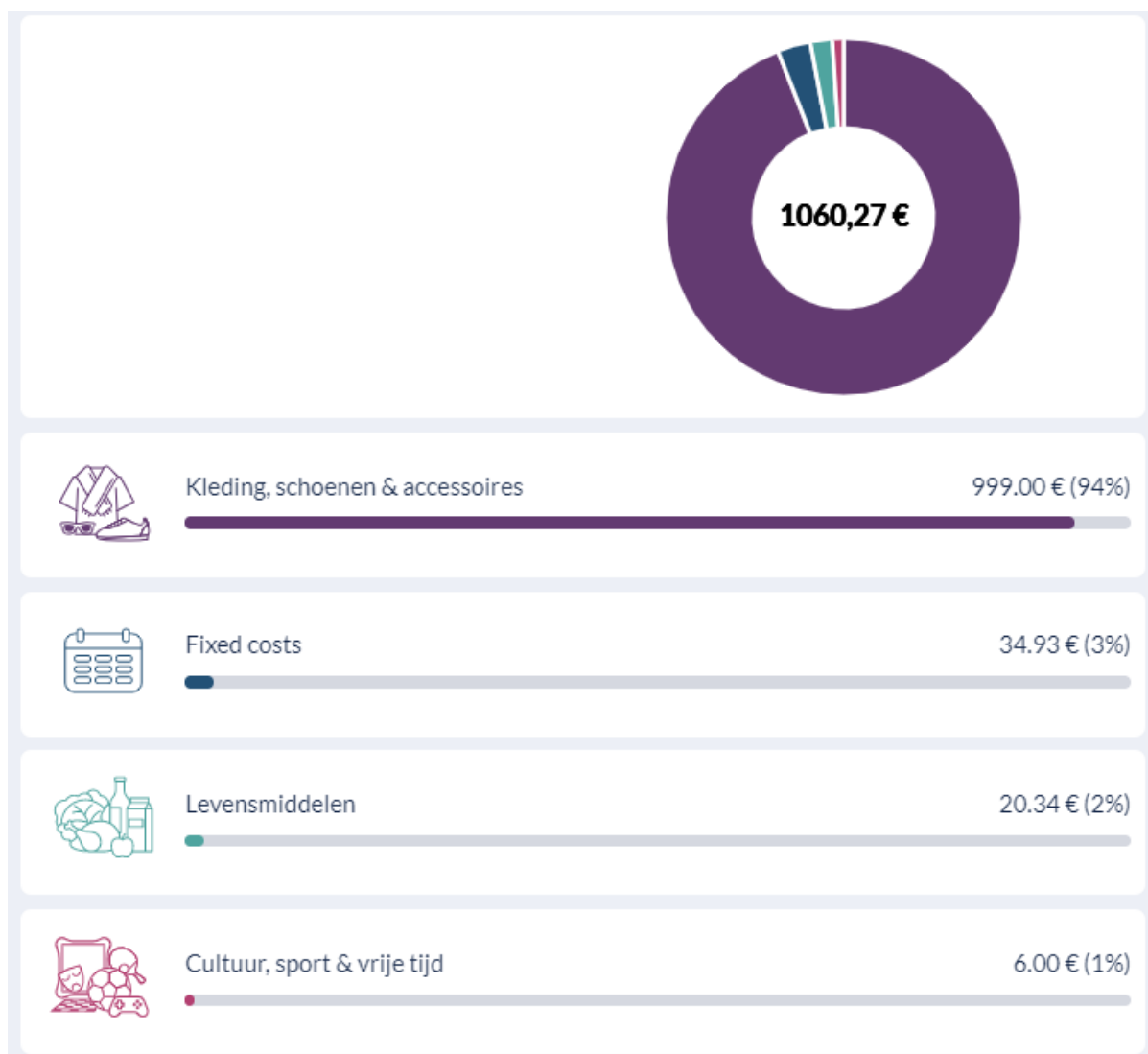
One of the HBS experts points out that the presence of an interviewer can be steering and motivating to pull together all the correct information. Working with a 'self-completion' questionnaire removes this. This should be accommodated by clear instructions such as short instructional videos or a 'wizard' that takes the respondent by the hand.

Expenses related to insurance taken out by the household could also be entered under fixed monthly expenses in the cost diary. Three respondents have a preference for including such expenses in the cost diary. The remaining three have no preference.

Answering the questions on in-kind benefits received by the household was assessed as straightforward by all respondents. Three respondents felt that it would be appropriate to include in-kind benefits in the cost diary in the MOTUS application (by analogy with monthly fixed costs).

Feedback on spending patterns

Through 'View your insights', respondents are shown a visual overview of all recorded expenses. The MOTUS application creates a pie chart so you can see at a glance how the costs are distributed. You will also see them neatly listed below that from large to small.

	Kleding, schoenen & accessoires	999.00 € (94%)
	Fixed costs	34.93 € (3%)
	Levensmiddelen	20.34 € (2%)
	Cultuur, sport & vrije tijd	6.00 € (1%)

The overview is well received. It is useful to find out what the most money was spent on in percentage terms. Towards the future, perhaps income could also be included in the dashboard so that respondents can check how their daily expenses compare with the available budget.

Unfortunately, the overview is currently not flawless. The way expenditure is allocated to a particular category is not clear. For example, a sports season ticket is largely allocated to 'Fixed costs' and only minimally to the 'Culture, sports and leisure' category. Some expenses are categorised completely incorrectly. The purchase of a car, for example, is assigned to the 'Clothing, shoes & accessories' category. As a respondent, you cannot adjust this manually. Such erroneous feedback can cause a big snap in respondents' confidence in the application.