ESSnet Smart Surveys

Grant Agreement Number: 899365 - 2019-DE-SmartStat

Link to our CROS website

Workpackage 2

Smart Survey Pilots

Deliverable 2.2 - WP2.2 (Time Use)

Version 2.0, 22-03-2022

Prepared by:

Patrick Lusyne (Statistics Belgium, Belgium)
Joeri Minnen (hbits/VUB, Belgium)

Workpackage Leader:

Barry Schouten (CBS, Netherlands)
Jg.schouten@cbs.nl
telephone : +31 70 3374905

Disclaimer: Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or Eurostat. Neither the European Union nor the granting authority can be held responsible for them.

Co-funded by
the European Union

SUMMARY

The modernization of the production of official statistics faces challenges related to technological developments, cost reductions, and growing privacy concerns. At the same time, there is a high need for shareable, scalable, and comparable data. Modularity is a key principle to address these challenges.

This paper considers three instances of modularity; subject, process, and methodological modularity. We map each of these manifestations of modularity to our experiences during two pilots that were part of the ESSnet on smart statistics. MOTUS, a data platform and smart survey application developed by the Free University Brussels and Hbits, was deployed to collect data in these small scale studies. MOTUS features functionality, deeply rooted in its architecture, that highly facilitates adopting a modular strategy to research.

1 Introduction.

Today, statistical offices face challenges and changes in producing official statistics [1]. Technological developments create the need for paradigm shifts in methodology. This is further fueled by budget costs, modern societal changes, and challenges that create new user demands for high-quality data and statistics on the one hand and fuel people's growing privacy concerns and suspicion towards official statistics [2]. Consequently, the modernization of the statistical business process model is characterized by a quest for scalable and comparable data collected in 'clever' ways, which are less costly for the statistical offices and less burdensome for the respondents.

Together with shareability, modularity is often put forth as a key principle to realize this objective. However, while the statistical community has invested a lot in conceptualizing shareability in the last few years, modularity is often implicitly assumed and commonly remains vague.

Work package 2 of the ESSnet on smart statistics allowed us to gain a lot of experience with MOTUS. MOTUS is a software platform developed by Free University Brussels to modernize time-use data collection. The acronym MOTUS stands for Modular Online Time Use Survey, explicitly highlighting the application's modular capacities. Experimenting with MOTUS in the ESSnet allowed us to have a hands-on experience with modularity rather than having theoretical consideration about it.

In this document, we report on this experience in various ways. Firstly, as MOTUS' modular functionalities are firmly rooted in its architecture, we briefly go into architectural-related elements that are key to that capacity. This information largely overlaps with other deliverables of the ESSnet on Smart Statistics and with the material in external sources. Therefore, it's a kind of concise recapitulation focusing on modularity.

We then map concrete elements of the two pilots we organized to several aspects of modularity. Modularity has been a real driver behind these pilots, and it even explains particularities regarding their design. The first pilot, aimed at functionality testing, consisted of several surveys concerning various subject matters. The second pilot was

on usability and mobility and comprised several surveys, a diary, sending notifications, and a geofence.

Modularity is a broad concept and, like we already mentioned, somewhat undefined. In this paper, we consider at least three instances of modularity, (1) thematic modularity, (2) process modularity, and (3) methodological modularity.

(1) Thematic modularity refers to a measurement instrument's ability to combine several themes or domains in one research flow in a modular way. Designing and organizing questionnaires using thematic modules is the most elementary example of this. This kind of modularity was at the heart of the reform of the Social Surveys in the ESS.

(2) Process modularity relates to the modularity of the production process. The GSPBM's conceptualization of the statistical production chain, implying breaking down the entire production flow into different subprocesses, is an explicit modular approach to the statistical production process [3].

(3) Finally, methodological modularity stands for the ability of an application or a platform to combine data collection methods in one research flow. This resembles a mixed-mode data collection, but a modular approach implies a design in which distinct data collection methods are used in different parts of the research flow by the same research subjects. The ability to combine survey and diary research is a basic example of this. However, a much more critical feature for smart statistics is the capacity to integrate active and passive measurement, survey and sensor data, for example.

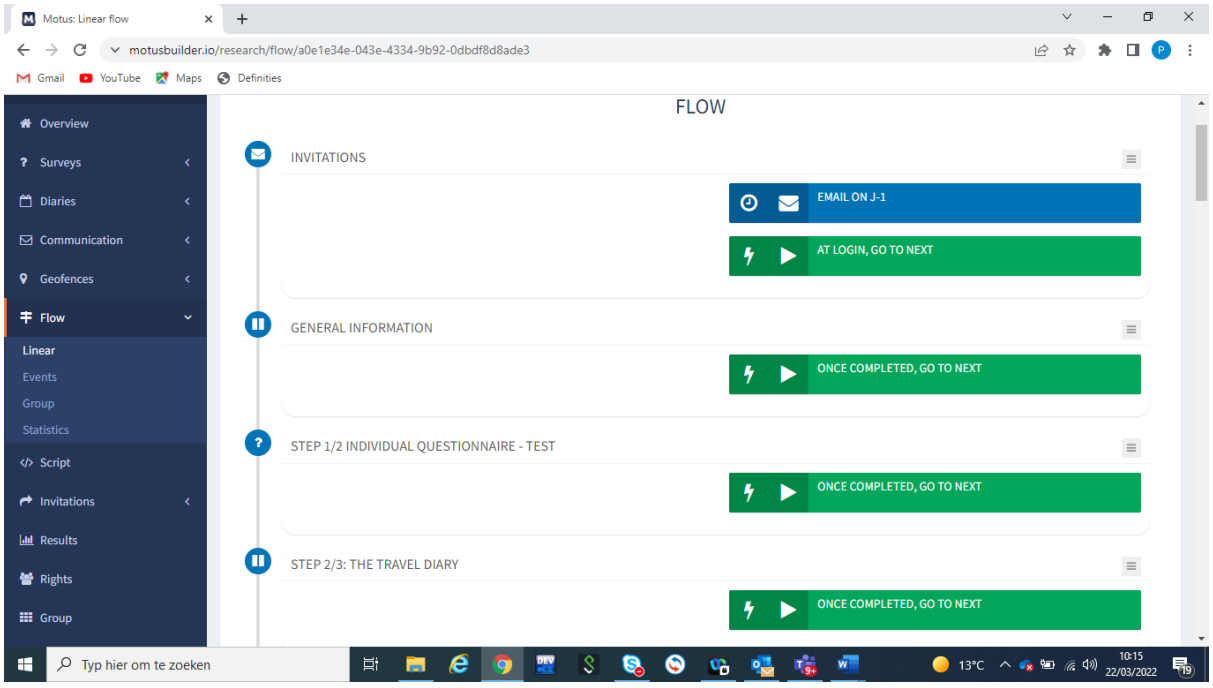
This elementary classification arose from our work in the two pilots. Therefore, we use it to map various activities and experiences in the pilots to modularity in general and its distinct instances in particular.

2 Modularity and MOTUS' architecture.

Perhaps the most perceptible manifestation of how deeply modularity is rooted in MOTUS is the design of its back-office, and particularly the so-called 'builders' that it features. Entirely in line with the CSPA logic, these builders are separate applications that integrate specific functionality. The functionality of each one of the builders is accessible via the overarching 'MOTUS builder'. The MOTUS builder operates in a GUI-like manner. It enables usage of the other builders essentially by clicking buttons.

A builder particularly rendering modularity, is the research builder. This builder serves to create the research design of a study. It enables the integration of all research steps in a research flow, chronologically or simultaneously. Figure 1 is an extract of the research flow of the second pilot we organized. This view shows the research builder as presented in the overarching MOTUS builder.

Figure 1. Extract of a research flow. A perspective from within the MOTUS builder.



This architecture with different builders each offering specific functionality, and the overarching MOTUS builder brings all three forms of modularity in reach.

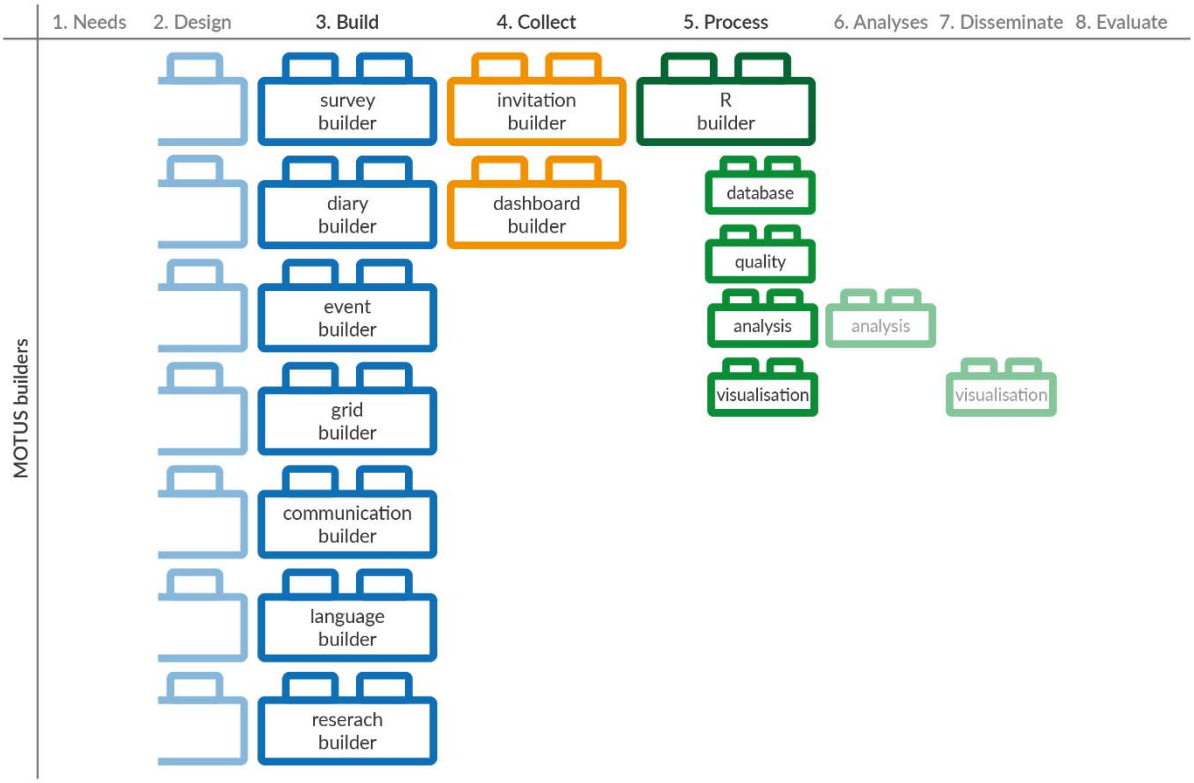
From the perspective of subject modularity (1), using the survey and research builder make it so easy to produce and integrate several domain-specific surveys that over-eager statisticians should take care not to exaggerate.

Regarding process modularity (2), the reach of the various builders makes that almost an entire GSPBM cycle can be covered without leaving the MOTUS environment. Figure 2 (next page) gives an overview of the main builders and how they relate to the phases of the GSPBM.

Finally (3), the development of a geolocation plugin or microservice in MOTUS' backend, deployable through the MOTUS builder, is an explicit example of methodological modularity. This plugin (partially still under development[1]) aims at collecting passive location data in a separate database to process and contextualize it with additional information from external databases such as OpenStreetMap. However, MOTUS' capacity to integrate this data with active data from surveys lifts this contextualization way beyond its reach when only using passive data, especially when passive and active data collection interact. This allows obtaining subjective (contextual) information, impossible to collect passively. For example, in the case of passenger mobility, it's essential to have an insight into the means of transportation that people use to move around, but also if there are alternatives and if people consider these.

[1] a geolocation plugin or microservice is being developed within the framework of a proof of concept (PoC) contract between Eurostat and Sogeti to further refine this functionality.

Figure 2. Overview of MOTUS builders and their main relation to the GSBPM.



3 Modularity and the pilots.

Although leading to a fundamentally different experience, the two pilots we organized shared several features. Both were small, and participants were recruited among NSI collaborators and colleagues. They also shared the overarching aim of gaining as much experience as possible with MOTUS within the 'safe' context of an experimental setting.

First pilot. The goal of the first was to assess MOTUS' functional dimension; if and how well the application performs from a technical point of view. To make this evaluation, the colleagues from the free University Brussels/ Hbits operationalized four questionnaires into MOTUS; a complete questionnaire on time use, an elaborated module on mobility (including sub-modules on public transport and cycling), and a smaller modules labor time. This study had a fairly elementary research flow. Participants were invited to enter the study (download MOTUS and log in) and to

respond to the various questionnaires. Bilateral exchanges between Statistics Belgium and the participating National Statistical Institutes served to assess MOTUS' functionality. The results show that MOTUS functions well. For a more detailed account and conclusions of this pilot, see deliverable 2.10.

Second pilot. The nature of the organization of the second pilot was different. We (Statistics Belgium) operationalized the research flow largely by ourselves in MOTUS, with support of the Free University Brussels where needed. While the focus was on usability and mobility from a substantive point of view, we also deliberately elaborated the research flow to stretch MOTUS' capacities to their edges. Table 1 gives an overview of the research steps in the study. This table comes from deliverable 2.1.

Table 1: Detailed overview of the research flow.

R1: Producing indicator passenger mobility

R2: Assessing Usability

R3: Gaining experience with MOTUS

[]	STEP/ELEMENT	TIMING	CONTENT	OBJECTIVE
0	Initializing communication recruitment	W-2	Invitations, recruitment, logins, briefing	R3
1	Mobility survey	W-1	Mobility	R1/R3
2	Usability survey	W-1	Specific usability issues survey	R2
3	Diary	W-0	Main diary – reference day – activity related questionnaire	R1/R3
4	Notifications	W-0	Parallel to main diary – 11AM – 3PM reference day	R2/R3
5	Geofence	W-0	Parallel to main diary – questionnaire	R2/R3
6	UX survey	W-0	Overall user experience survey	R2

A critical element in this flow was that we 'tweaked' the diary to make it more suitable to study mobility. This adaptation resulted in a lightweight diary in which principally trips were registered and only minimal information on other activities.

Nevertheless, the overall complexity of the research flow still made the fieldwork of the pilot quite challenging, which made the learning experience all the richer. Deliverable 2.1 gives an elaborated overview of the results of the mobility study, the evaluation of MOTUS' usability, and an account of our experiences.

Thematic Modularity. To the question which instance of modularity is the most important one, many statisticians involved in fieldwork probably would answer thematic modularity, getting rid of the stovepipes and being able to organize modular surveys for several domains within one ecosystem.

Thematic modularity was also for us a prime driver, both for the content and the set up of the pilots. Implementing the modular nature of the first pilot was very simple. The integration of various questionnaires across different domains proved to be straightforward.

In the second pilot, we took things a step further, however. By tweaking the diary, as explained above, we were able to calculate basic but nonetheless critical parameters to the field of passenger mobility statistics. Imperative to the Green Deal, passenger mobility as a research strand has gained substantial weight, and the need for reliable and comparable parameters has grown accordingly.

Although our study was tiny and experimental, it proves, in our vision, that an application like MOTUS can relatively easily be adapted to serve other domains than its native Time Use environment. Moreover, MOTUS' platform features and the ease with which statisticians can harness these capacities in a fully autonomous way make that there's little to obstruct a widespread use of this technology, apart from (national) legal or organizational specific constraints.

Process Modularity. Annex 1 gives a detailed overview of MOTUS' coverage of the GSPBM production cycle. The core function of MOTUS relates to the building, collecting, and processing phases of the production chain. The second study was again the one that led to the most tangible, hands-on experiences with these different phases, or at least with the building and collecting phase. The analysis phase was, by choice, largely done outside of MOTUS (in SAS). We consider the freedom to leave the MOTUS environment also as an expression of the platform's modular nature.

Applying the logic of the GSPBM production chain in MOTUS is quite simple. Building and integrating measurement instruments using the designated builders take some craft, but the associated learning curve is limited, as we experienced during the pilot. On the other hand, managing the fieldwork of mobile research turned out to be the real challenge, even of a study with a small scale like ours. However, we also experienced first-hand that MOTUS features functionality, including direct communication with the participants and follow-up, that helps get the job done.

Methodological Modularity. Because of the context of the Essnet on smart statistics, the importance of combining survey research with a lightweight diary may have been somewhat underplayed in the various deliverables we already produced. However, while combining survey and diary research may not be called a smart feature, it may have relevance for many research strands. Researchers tend to avoid this combination as it's challenging to operationalize and increases response burden significantly. Nevertheless, combining a lightweight targeted diary in the context of a smart survey instrument might alter this situation and facilitate the use of diaries outside the traditional time use perimeter.

Additionally, as diaries imply by definition the registration of time and activities, diary research shifts the research focus, much more than regular surveys, to the interdependence between time and context. 'Mobile diaries', in particular, may decrease the span between an activity and its registration significantly, and questions on context can promptly be presented. This capacity increases the reliability of the registrations, alleviating the impact of recall bias, for example. More importantly, however, it also allows getting data that is otherwise hard to collect, like short-lived emotions and people's spontaneous motives.

The advantage of having a short activity-registration span remains purely theoretical, though, if respondents miss the reflex to register an activity as soon as they end it. Consequently, while collecting data with a mobile diary is clever, it needs to be made smart by harnessing the technology's capacity to prompt participants to register their activities immediately.

We send notifications at fixed moments throughout the diary reference day in the second pilot. However, this can't be labeled smart as it only partially considers the context (the hour of the day). The geofence is a genuine smart feature, on the other

hand. It passively traces participants to prompt them to respond to questions or register activities based on their location. Geofences have huge potential, in our opinion. While tracing persons to get insight into entire itineraries is only valuable for mobility research, geofences may have much wider use, revolutionizing all kinds of research strands.

However, modularity is a critical condition to this. Wider use of geolocation and tracing functionality, of any smart feature, is only of value if it's embedded in an architecture allowing modular integration with other data collection methods. However small the scale of the pilots, and notwithstanding the many issues we experienced, the MOTUS platform clearly features this capacity.

4 Conclusions.

Modularity isn't a uni-dimensional concept; in this paper, we distinguished between three manifestations of it. Thematic modularity refers to the capacity of an application to combine the collection of data for several research domains in a modular way. Process modularity is about modularizing the production chain. Finally, what we call methodological modularity, relates to the ability of an application or a platform to combine passive and active measurement.

Experiencing modularity first-hand was a prime driver behind the two pilots we organized with MOTUS in the Essnet on smart statistics. Many elements in these pilots were introduced with modularity in mind. Consequently, it was quite straightforward to map various experiences during the pilots to one of the three instances of modularity. Although these pilots were small, they prove, in our opinion, that the reference to modularity in the acronym MOTUS is just.

The principle justification for that claim is found in MOTUS' back- and frontend. Modularity is at the very heart of MOTUS' architecture. The so-called builders ar the most manifest expression of this. Each of these covers a vital piece of functionality, and all are integrated into one overarching work environment, called the MOTUS builder. This integration and the accessibility of the MOTUS builder highly increases the usability and deployability of MOTUS from a statistician's point of view.

The latter is essential. Modularity, in whatever manifestation, shouldn't be a goal as such. Its actual value lies in its connection to other features, such as flexibility and usability of an application, and shareability. This paper isn't, therefore, a plea to embark on the ultimate modular tool quest. What we need are workable applications, no matter how modular they are. However, our experience with MOTUS, explicitly by harnessing the application's modular capacities, taught us that modularity and usability necessarily go hand in hand.

Annex 1: Applying GSBPM to MOTUS.

As already mentioned, the core function of MOTUS relates to the building, collecting, and processing phases of the production of time-use statistics. Given that these phases are prompted by information from previous phases, we will briefly mention the other phases as well.

DESIGN PHASE.

The design phase leads to the definition of the statistical outputs, concepts, methodologies, collection instruments and operational processes. When it comes to the collection of official time-use statistics in Europe, the design phases will be largely informed by the HETUS guidelines [4], but other national and international standards may apply outside Europe. These guidelines and standards, for example, provide input for metadata that are part of the design output (2.1) and variable descriptions (2.2) of which the Activity Classification List (ACL) is an important element. When it comes to TUS, perhaps the most important subphase of the design includes decisions about the data collection method (2.3). At this point, the GSBPM must be understood as a matrix with many possible, iterative, and interdependent paths. Certain design decisions might be informed by the builders currently available in MOTUS to the same extent that the build phase can be informed by the need to design and add builders to the existing arsenal of MOTUS. The main collection instruments and methods to be designed for TUS are questionnaires (e.g., individual, household, activity or context questionnaires), and a time-diary. For an online TUS, at least some instruments are supportive of the collection, such as online communications, mode of collection and different language settings in MOTUS.

In contrast to traditional surveys, such as the Labour Force Survey (LFS), which are composed of one-off questionnaires a TUS includes multiple research stages, such as the household questionnaire, the individual questionnaire, and the time-diaries. Therefore, the design collection subphase for a TUS should also include the design of the research flow, that is, all the research stages a respondent must go through from the moment the respondent logs on to MOTUS. This also includes decisions on the conditions under which the respondent is allowed to proceed to a next stage, the conditions under which inactive respondents are reminded of their tasks at hand, and the conditions under which inactive respondents are drop-outs. Conditions to proceed

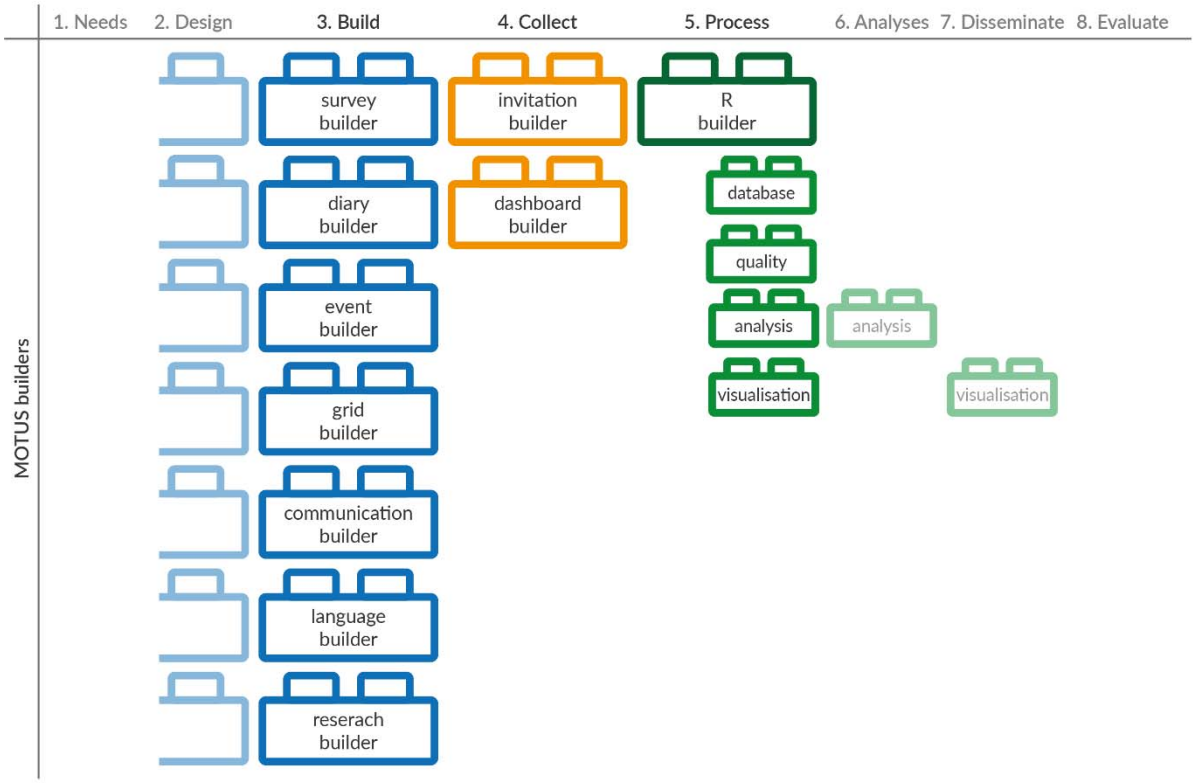
to a next stage might depend on completion of a stage by the respondent or by all group members (e.g., when synchronising time-diaries at the household level) or on passing a certain date (e.g., when randomly assigning time-diary days). Conditions to remind or excluded respondents might depend on the number of inactive hours or days passed after entering a certain stage.

For national time-use surveys, NSIs typically use the national population register as the frame for drawing probability samples of individuals and include the households of these individuals in the sample (2.4). From here the invitation strategy for traditional paper-and-pencil TUS was straightforward. Interviewers would visit the sampled household, conduct the household and individual questionnaires, and leave the time-diaries after explaining how they work and on what days all household members from a certain lower age limit must complete the diary. In this example, the individual participants to the time-use survey are clustered within a household and HETUS guidelines require all individuals within a household to record the same diary days. More generally, if individuals are clustered within groups (e.g., members of households, colleagues of departments, students in classes, and so on) and if synchronous time registration is expedient, at least part of the workflow design (2.5) involves mapping a strategy to invite group members and synchronise their time-diary registration. Multiple strategies are conceivable, such as inviting a reference person that compiles the group, prefilling groups based on other administrative data, or inviting all group members individually.

BUILD PHASE.

Decision made in the design phase inform the build phase. In this phase, the research components and research flow are built and implemented in MOTUS. Building and implementing components and research flow is done using so-called MOTUS builders. Currently, MOTUS contains seven builders that support the build phase of the GSBPM (see Figure 2'). These builders are the *survey builder*, the *diary builder*, the *event builder*, the *grid builder*, the *communication builder*, the *language builder*, and the *research builder*. These builders are to be seen as building blocks that can be stacked onto each other without priority.

Figure 2'. (repetition fig. 2) Overview of MOTUS builders and their main relation to the GSBPM.



As the names suggest, all individual builders serve a specific purpose (e.g., creating a survey, adding an activity list to a time-diary, writing communication, …). However, at the same time they are related to each other. For example, time-use surveys generally gauge the context of activities that are recorded in the time-diary (e.g., with whom and where the activity was undertaken, to what extent the activity was satisfying, if the activity involved the use of information and communication technology, …). These short context questionnaires are created in the survey builder and attached to activities in the ACL in the diary builder. Similarly, the language builder defines the languages in which the study is offered. The language builder not only translates all hard-coded text elements such as buttons and menus but also makes a translation environment per defined language available in all other builders to enter translations.

Survey builder. The survey builder allows to create questionnaires based on over 20 question types. In the survey builder, answer-based routing can be created and given answers can be reused in the same questionnaire or in another questionnaire further

down the research flow. For example, if someone is interested in distinguishing childcare by each of the children present in the household, one questionnaire could inquire the names of the children and another questionnaire could present these names as a context question when someone registers childcare in the time-diary. The survey builder contains a repository with previous questionnaires and questions for reuse.

A typical TUS would use the survey builder to construct a household questionnaire, an individual questionnaire and context questionnaires. As mentioned above, context questionnaires are linked to activities in the ACL to capture the context of these activities. The HETUS guidelines translate to two context questionnaires that are basically the same and only differ in the case of a travel activity. The question about where the activity took place, is then replaced by a question on the mode of transport. Theoretically, though, a different context questionnaire can be attached to each different activity in the ACL, once defined in the survey builder.

Diary builder. The diary builder foresees in the setup of the time-diary. At the core of the time-diary is a list of activities that respondents use to register their daily life. Paper-and-pencil time-diaries typically require respondents to write down activities in their own wordings. The HETUS guidelines, then, provide instructions and an activity coding list for post-coders to code these entries. This activity classification list (ACL) follows a three-level tree structure. On the main level, activities collapse into categories or domains of life that people recognise as such in contemporary societies. They are: (0) Personal care, (1) Employment, (2) Study, (3) Household and family care, (4) Voluntary work and meetings, (5) Social life and entertainment, (6) Sports and outdoor activities, (7) Hobbies, (8) Mass media, and (9) Travel and unspecified time use.

In MOTUS the HETUS based ACL is translated into Online Activity Classification List (OACL) in the same three-level structure. The OACL omits the high costs of post coding verbatim diary entries.[2] The diary builder, however, facilitates the creation of customized activity lists using one up to three levels and as many activities or activity

[2] Costs for verbatim diary entries are even higher because most NSIs use multiple coders per diary to reduce inter-coder variability that might results from coders' interpretation or the readability of the entries.

categories in any given level as needed. Each activity in the list has a unique corresponding code for data processing. Currently, two (different) activity lists can be used in the time-diary design, one for the main activity and one for the secondary activity. (Note that recording secondary activities is part of the HETUS guidelines.)

For example, the microservice of geolocations could suggest travel or work activities if the respondent is at a geolocation that is defined as the workplace.[3]

Next to the activity list, the diary builder also allows to set all the parameters of the time-diary. The diary builder distinguishes between the diary period and the registration of so-called focus days. The diary period is the time the diary is accessible to the respondent. The focus days are the days the respondent is requested to complete (within a diary period).

Event builder. The event builder supports the inclusion of events in a study. Events follow the if-this-than-that (ITTT) approach and, thus, are triggers that are pulled if a certain condition is met. These conditions and the action they initiate are defined in the event builder. Most well-known is the experience sampling method (ESM of beeper method) that ask respondents to answer questions on random or pre-programmed moments over a certain period. MOTUS expands this method by coupling sensor data to trigger events based on so-called microservices that use an API to communicate with MOTUS. These events can on the one hand ask the respondent to perform a specific action (e.g., answering a short questionnaire) and on the other hand show tentative entries in the respondent's diary, which can reduce the registration burden and increase the quality of the registration.

For example, MOTUS currently includes the geolocation microservice. This microservice allows for the collection of geolocation points. These points can be enriched with publicly available data (e.g., OpenStreetMap) or personal data (e.g., respondent identifying a location as home or work) and delimited with a radius around these points (so-called geofences). Each geofence has three states, that is, enter, exit, or dwell. Any of these states can trigger an event such as sending a questionnaire of

[3] In accordance with GDPR guidelines, respondents give consent to these suggestions by activating these supportive aids. Each suggestion will be presented as tentative data that need to be actively confirmed, edited, or rejected by the respondents.

activating an algorithm that, based on context derived from locations databases (e.g., Google Places, Foursquare Places or OpenStreetMap Overpass)[4], provides tentative data to the respondent. Subject to respondents' consent to track movements, respondents view tentative data from the geolocation microservice in their MOTUS time-diary where they can confirm, modify, or reject this tentative data.

In the future, microservices can be developed and added for any sensor data. For example, sensors in health such as heart rate monitors or sleep monitors, or sensors in environmental science such as CO2 meters or temperature meters. This allows MOTUS to further contextualize people's daily lives with an interdisciplinary lens. Of course, the respondent will be asked for explicit consent for each sensor that is used. Moreover, events – or the use of microservices – only become effective when designed as part of a research state in the research flow (see research builder). Both steps ensure that sensor data is captured only in function of the study and with the consent of the respondent to protect the privacy of respondents.

Grid builder. The grid builder is used when the unit of participation is not the individual level, and when the participation of group members needs coordination. The grid builder departs from a reference person who creates a group by adding members, specifying all relations between members, and answering questions about the characteristics of all members. The reference person might also complete a group level questionnaire (e.g., characteristics of the group). These questionnaires are prepared in the survey builder. Based on the input to the grid, group members can be invited to take part in a study. Invitations might be subject to the reference person's consent or restrictions based on group member characteristics (e.g., age). These restrictions are defined beforehand. If the reference person provides group members' email addresses, all group members who are eligible to participate will receive an invitation via email with their initially assigned personal credentials, which can be altered once participation in a study. Optionally, a copy can be sent to the reference person. Hereafter, the next step of each invited group member is to confirm and/or update their e-mail address and change the initial password. Both actions are mandatory to protect the privacy of each respondent.

[4] In line with GDPR regulations MOTUS anonymously requests data from these services.

For example, a TUS that follows the HETUS guidelines is a household survey. The group, in this case, is defined as all household members living under the same roof and all household members aged 10 years and above are invited to participate. The reference person of a household composes the grid for the household by adding household members, specifying relationships (e.g., mother-daughter, partners, siblings …), and answering questions about household members (e.g., about care for young children). The reference person also completes a household questionnaire (e.g., about household help, household income …). Hereafter, all eligible household members are invited to complete an individual questionnaire, a time-diary with two focus days and end-of-day questionnaires.

Communication builder. The communication with the respondents is another core element of conducting an online survey, especially a TUS where the survey exists of multiple stages (e.g., the invitation, a household questionnaire, an individual questionnaire, and a time-diary). The absence of an interviewer leaves at least four ways of communicating throughout the data collection process: email, text messages, push messages, and static webpages. All communication is defined in the communication builder.

The communication builder also allows for the import and selection of email templates, personalize the contact information of a specific study (i.e., name, email, phone, postal address …), and add customized static webpages, for example, with FAQs or the study's privacy declaration, and with dynamic URLs to reset a password or to activate a new MOTUS user account.

Language builder. The language builder makes it possible to define multiple languages for each study, allowing respondents to choose the language in which they want to conduct the study. Languages become available in the language builder after they have programmed in MOTUS. Indeed, first, all system elements (e.g., menu options, labels of fields, button text, …) must be translated in the source code via translation files. Currently, English, French, German, and Dutch are available in MOTUS. Once available, languages can be activated, and a default language must be chosen. The translation of collection instruments is integrated in the respective builders. All elements in each of the builders are created in the default language and

translation fields are automatically provided for each language specified in the language builder.

Offering a study in different languages is done for practical reasons (e.g., because a country has several national languages, such as Belgium or Switzerland) and to increase the ease of use for the respondent. Which language the respondent sees depends on three conditions or language priorities. If the preferred language of the respondent is known before the start of the study, the preferred language can be specified at the time respondents are added to the study. MOTUS will be presented to the respondent in the preferred language. If no language preference is known, the language in MOTUS is based on the device language settings. If none of the above settings is decisive, and/or none of these languages are available, the default language is used. Once respondents' set or change their language preference, it will be remembered on all devices used to participate in the study.

Research builder. All the built collection instruments (e.g., the questionnaires, the OACL, the diary, and events) and communication are stored in their respective builders. The research builder constructs the collection process or so-called research flow (i.e., subphase 3.2 of the GSBPM). The research flow brings together all collection instruments and communication and specifies the conditions under which respondents go through the collection process. The research flow is linearly oriented, but events that are added can be recurring. This linearity follows from the different states that respondents must go through to successfully participate in a study. These states typically exist of collection instruments (e.g., filling out a questionnaire or keeping a time-diary). Moving through states are called actions and these actions are conditional. The conditions are defined in the research builder (e.g., when completed the questionnaire go to the time-diary). Within a state, different conditions can result in different actions (e.g., when questionnaire is inactive for several days terminate the respondents' participation). Actions are communicated to the respondent based on the means of communication that are created in the communication builder (e.g., when respondents complete a state, a web page is shown and/or an email is sent with instructions for the next state). Additionally, communication criteria can be defined based on the progress within a state. These criteria are typically time based (e.g., a reminder is sent when a respondent's time-dairy is inactive for 24 hours, or an encouragement is sent when a respondent recorded three consecutive days in the

time-diary). Finally, events are attached to a state as a specific action. The event is defined in the event builder and only applies to the state it is attached to. For example, the use of location tracking can be defined so that it is only applicable during the time-diary stage. This way, the event is not only responsive to the research question but also to the privacy of the respondents.

For a TUS that follows the HETUS-guidelines, the research flow exists of several states, such as the completion of a household questionnaire and grid composition by the reference person, the completion of individual questionnaires by all household members eligible to participate, and the completion of two focus days in the time-diary. New users of MOTUS will go to additional states of activating their MOTUS user account (e.g., validating their email address), mandatorily changing their initial password, and completing their user profile.

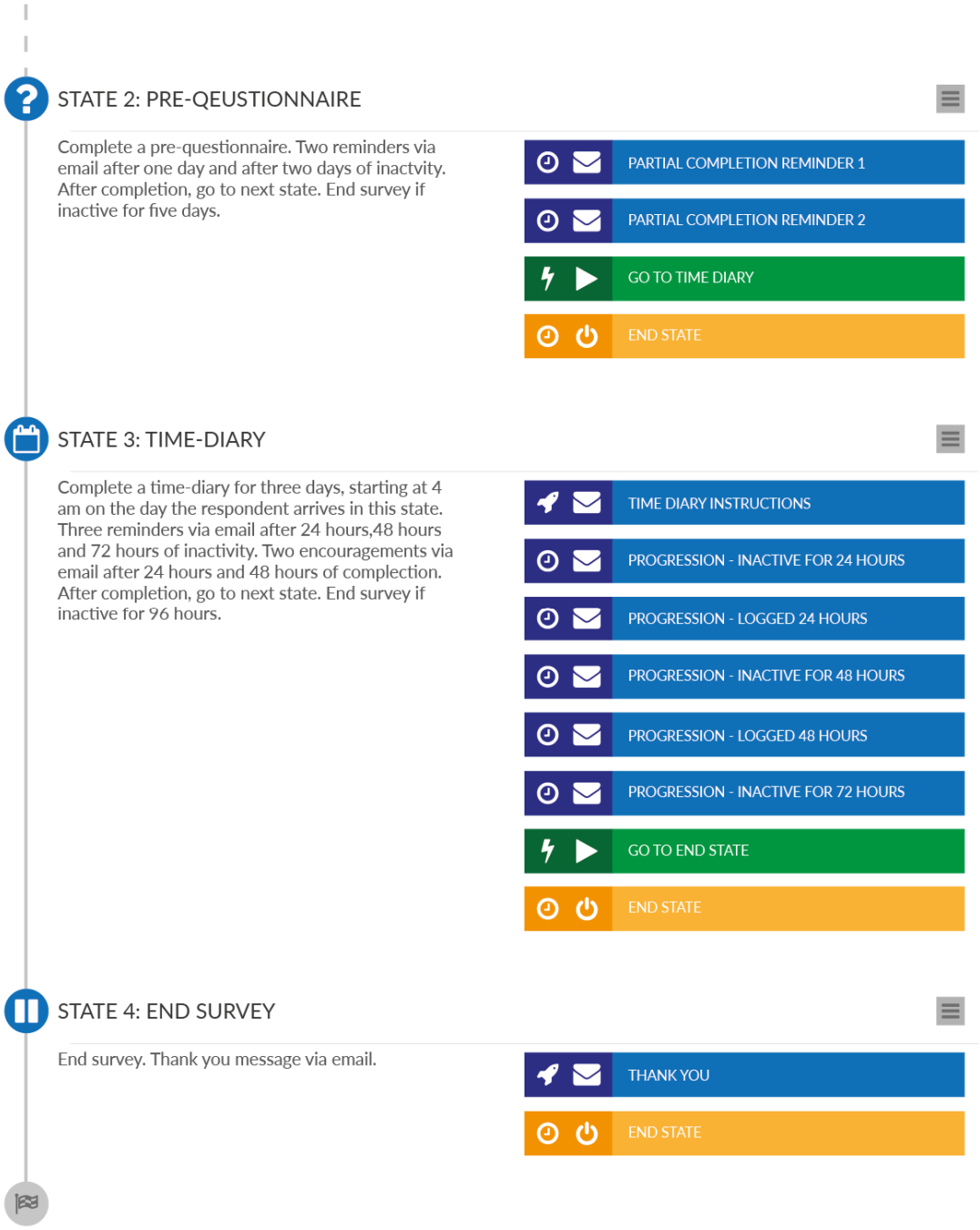
Figure 3 gives an example of a simplified research flow of a TUS that involves an individual pre-questionnaire and a three-day time-diary only from respondents that already activated their MOTUS account, changed their password, and completed their user profile.[5] The states are listed on the left-hand side and the actions on the right-hand side. There are three conditions that trigger actions: automatically when entering the state (indicated by the rocket ship), time-based (indicated by the clock), and automatically when completing the state (indicated by the lighting bold). There are also three types of actions: actions that involve communication (in blue), actions that move respondents through the linear flow (in green), and actions that terminate or finalize the respondents' participation (in orange).

The research flow in Figure 3 then shows that respondents that are in the state of completing the pre-questionnaire will receive up to two time-based reminders to complete the questionnaire. They will automatically go to the time-diary state once they have completed the questionnaire, or their participation will be terminated when they remain inactive. Once respondents enter the time-diary state they will automatically receive an email with instructions. In this state there are five time-based actions that include communication: three reminders (i.e., after 24, 48 and 72 hours of

[5] Currently, MOTUS experiments with the principle of reverse engineering. This involves first building the research flow based on identified needs (i.e., phase 1 of the GSBPM) and work backwards from the research flow to design and build all relevant (collection) instruments (i.e., phase 2 and 3 of the GSBPM).

inactivity) and two encouragements (i.e., after logging 24 and 48 hours). Again, respondents that complete this state automatically go to the end-of-survey state and respondents that remain inactive will see their participation terminated. Once in the end-of-survey state, respondents automatically receive an email that thanks them for their contribution. The TUS is then completed and closed.

Figure 3. Simplified research flow of a TUS that involves an individual pre-questionnaire and a three-day time-diary only from respondents that already activated their MOTUS account.



Other analyses and dissemination components (i.e., subphases 3.2 and 3.3 of the GSBPM) are supported by the R builder, which is further discussed under the process phase (i.e., phase 5 of the GSBPM). The complete survey, that is, the research flow with all its collection instruments, communication and language support is accessible via the so-called MOTUS builder environment. This environment can be reached via the URL www.motusbuilder.io and is available in four different modes. The survey is built and tested (i.e., subphase 3.5 of the GSBPM) in the 'development mode'. The 'active mode' starts the collection (i.e., subphase 4.3 of the GSBPM), and the 'pause mode' allows for temporary suspension of the collection process. The 'archived mode' stores the complete survey in the MOTUS repository for (re)use (i.e., conform subphase 3.1 of the GSBPM). Furthermore, the MOTUS builder environment is sharable through different architectures: through a data collection architecture, through a natively installed architecture, or through container-based virtualization.

COLLECT PHASE.

Once a complete study has been built, the collection process can be started. Before running the collection, the collection frame needs to be created and the sample needs to be selected (i.e., subphase 4.1 of the GSBPM). This subphase is supported by the invitation builder.

Invitation builder. The invitation strategy feeds respondents into the research flow. Depending on the designed frame and sample (i.e., subphases 2.4 of the GSBPM), the invitation builder allows to set up different types of inflow of respondents. Firstly, if an organisation external to MOTUS sends the initial invitation, the invitation builder allows to generate a list of unique usernames and passwords that can be printed in the invitation letters. This is most common in official TUS. The invitation builder is also able to generate usernames and/or passwords randomly or based on information that is fed into MOTUS.

Thirdly, if a list of potential respondents and their email addresses is readily available, this list can be uploaded in the invitation builder and MOTUS will send invitations directly to all potential respondents. For example, in the TUS to measure the workload of university professors of the Vrije Universiteit Brussel, email addresses were known and MOTUS invited respondents directly via email (Verbeylen et al. 2016).

When the survey is set to 'active mode' (i.e., subphase 4.2 of the GSBPM) and data collection is ongoing (i.e., subphase 4.3 of the GSBPM), the dashboard builder provides tools to follow up the survey's progress.

Dashboard builder. The dashboard builder provides insights in the progress of the collection. It allows to monitor response rates, the inflow and (moment of) dropout, time spent in and progress through the different states of the research flow, and quality of the imputed data (see Process, analyse, and disseminate). Decisions regarding the adjustment of the fieldwork, for example the need to oversampling certain groups or regions, can be made based on the information in the dashboard.

PROCESS, ANALYSE and DISSEMINATE PHASES.

The process phase (i.e., phase 5 of the GSBPM) processes the data and prepares the data for analyses and dissemination in later phases (i.e., phase 6 and 7 of the GSBPM). All phases exist of several subphases that are supported by the R builder. This builder links the R language and environment for statistical computing and graphics to MOTUS. The R builder contains the motusr package by default, but inclusion of other packages is possible. The motusr package exists of four smaller components (see Figure 2'), that support different subphases of the process phase of the GSBPM.

Database. Firstly, the motusr package integrates the data from all different collection instruments based on a Universally Unique IDentifier (UUID) (i.e., subphase 5.1 of the GSBPM) and classifies and codes the response categories of the various collection instruments (i.e., subphase 5.2 of the GSBPM). Additionally, it provides data fames per stage, such as a questionnaire or time-diary, para data, such as when and how often respondents' recorded activities in their time-diary, and META data, such as time and date of data creation or data quality (see further). Secondly, the motusr package allows data integration beforehand, for example, when applying the Only-Once Principle (OOP) to data provision. In the case of an official TUS, respondents are often drawn from the national registry. The national registry contains data such as date of birth and sex, which can be prefilled in respondents' profile or individual questionnaires via the motusr package. Thirdly, the motusr package facilitates the export of the finalized database in almost any possible format.[6]

Quality. The motusr package contains several quality criteria against which data can be checked, flagged, and edited (i.e., subphase 5.3 and 5.4 of the GSBPM). Quality

[6] This includes the most commonly used database formats such as .sav for SPSS, .dta for STATA, .Rdata for R as well as .csv, .xls, .xlsx and .xlm).

criteria can be defined based on several parameters, such as undefined time per day, the number of activities, the number of different activities, the number of activities rounded at the hour, the inclusion of displacements if the spatial context of activities changes, the presence of certain activities (e.g., sleep, eating or drinking). Additional quality criteria can be defined (e.g., no more than 1 hour of unspecified time per day or at least one sleep activity every 48 hours). Based on the number of quality criteria included and the limits set, time diaries are flagged in three different colours: green for inclusion in the final database, red for exclusion of final the database, and grey for manual evaluation and editing. Certain quality criteria can also be defined to provide feedback to the dashboard builder where the data collection is monitored (see Collect).

REFERENCES.

[1] Radermacher, Walter J. 2020. *Official Statistics 4.0. Verified Facts for People in the 21st Century*. Cham, Switzerland: Springer.

[2] EUROSTAT. 2018. *Power from Statistics: Data, Information and Knowledge – Outlook Report - 2018 Edition*. Luxembourg: Publication Office of the European Union.

[3] Salgado, D., e.a., (2018) *Data organisation and Process Design Based on Functional Modularity for a Standard Production Process.* Journal of Official Statistics 34 (4):811-833.

[4] EUROSTAT. 2020. *Harmonised European Time Use Surveys (HETUS) – 2018 guidelines – Re-edition*. Luxembourg: Publications Office of the European Union.

[5] Verbeylen, Julie, J Minnen, Theun Pieter van Tienoven, and Ignace Glorieux. 2016. *Taakbelasting en taakinvulling van het ZAP aan de Vrije Universiteit Brussel.* Reserach Group TOR, Sociology Department, Vrije Universiteit Brussel (Brussels).