WP2 Smart survey pilots

Intermediate meeting

Oct 7, 2020

DESTATIS Statistisches Bundesamt | INE Instituto Nacional de Estadística | Insee | Istat | Office for National Statistics | SCB | Statistisk sentralbyrå Statistics Norway | STATEC | Statistics Poland

ESSNET SMART SURVEYS

STATBEL België in cijfers

Outline WP2

- General update (5 min)
- Updates pilot studies (WP2.1 to WP2.4)(30 min)
- Demo/presentation time use subthemes (25 min)
- Break (5 min)
- Discussion legal-ethical-policy requirements (45 min)
- Next steps WP2.1 and WP2.2 (45 min)

GENERAL UPDATE WP2

ESSnet objectives for WP2

1. Identification and collection of the functional and technical characteristics of smart survey solutions;
2. Establishment and execution of a solutions evaluation protocol for existing data collection applications;
3. Developing pilot projects demonstrating the use of scalable secure private computing solutions to process individual citizen data without concentrating personal data at a single entity, in combination with appropriate technological solutions to provide full auditability;

HBS and TUS are explicitly mentioned
Citizen science/smart statistics feel and features

Deliverables

1. Report per pilot in terms of statistics and comparability to regular surveys
2. Report for Consumption and Time use on expansions of tools
3. Report evaluating respondent feedback (pilots 1 and 2).
4. Report evaluating interviewer assistance (pilots 1, 3 and 4)
5. Report on shareability across country-specific settings (all pilots)
6. Report on modularity and answering behaviour (pilots 1 and 2)
7. Template functional and technical specifications for WP3
8. Report on infrastructure and logistics specifications as input for WP3
9. Open source code, if applicable, and manual for tools deployed in pilots
10. Files in English and French for translation of app texts and search libraries
11. Overall final report including recommendations and look at future)

Deliverable 12/2.7 –specifications WP2 for WP3

Descriptions of pilots:

1. CSPA conceptual, physical, logical levels (WP3 subtasks 3.1.2, 3.1.4, 3.1.6)
2. Methodology level (WP3 subtasks 3.1.1, 3.1.5)
3. Logistics level (WP3 subtasks 3.1.3, 3.1.5)
4. Legal-ethical-policy level (WP3 subtask 3.1.4)

Levels are split in frontend and backend and crossed against {generic, country-specific majority, country-specific minority}

Mapping to smart features and to WP3 proof-of-concept elements.

Timeline in general

- Intermediate meetings online: October 6 and 7
- Tentative opt-in data phase 3 (WP2.1 – 2.2): Dec 1, 2020
- Tentative intermediate meeting in-person: February 2021
- NTTS session: Q1 2021
- ESRA/ISI sessions: Q2-Q3 2021
- End of phase 3 data collection (WP2.1 – 2.2): Jan 1, 2022
- Tentative final workshop ESSnet: Q2 2022

WP2 PILOT UPDATES

Motivation for themes and pilots

- Consumption:
  - Need for reduction of respondent burden;
  - Cross-country comparability;
  - Existing innovation (HBS);
- Time use:
  - Lends itself very well for modular approach;
  - Existing innovation (HETUS);
- Health:
  - New tools: Wearables;
  - Link to EHIS;
- Living conditions:
  - Well suited for citizen science features;
  - New tools: Indoor climate systems;
  - Link to EU-SILC;

WP2.1 CONSUMPTION

WP2.1 – Consumption

Current status:

- HBS app: 10 country versions, 6 new versions in ESSnet
- Receipt scanning part is being tested in NL
- Small interviewer field test in NL; Sept/Oct 2020
- Phase 2/2': Basic tests in WP2.1 countries on-going
- Phase 3 fieldtest: In NL planned for March/April 2021

WP2.1 – Country versions

What is configurable?

- UI text translations (through English demo version)
- NSI logo
- Product lists linked to COICOP classification
- Shop lists
- Language for machine learning computer vision of receipts
- EAN/GTIN product descriptions linked to COICOP

WP2.1 – HBS ‘corona-proof’ interviewer test

- 50 adresses, 5 interviewers
- Announcement letter with €5, followed by interviewer visit
- If Yes, log-in codes were given
- Motivation call by interviewer
- 2 weeks app use, incentive €20
- No recurring expenses, only app use
- Fieldwork runs until end of October

WP2.1 – Preliminary findings interviewer test

- Up to Oct 5: 14 participating households (±30%)
- Age matters
- Motivation call 50% of the cases answered
- If answered, no technical issues, only survey related
- Relationship with respondent
- QR code

- Overall; everyone (I & R) like the use of the app

WP2.1 – Country phase 2/2' tests

- Basic tests are on-going in WP2.1 countries
- Dynamic FAQ documents maintained for HBS app frontend and backend
- Evaluations split in three:
  1. General usability: 13 questions about UX, product search, data entry
  2. Country-specific frontend: 11 questions about content, data models, product/shop lists
  3. Country-specific backend: 13 questions about data collection monitoring, data access, need for other modes

Results expected before end of 2020.

WP2.2 TIME USE

Work package 2.2 Time Use

- Detail on legal issues, original set up, (near) future implementation, partial redesign of scope
- Legal issues: DPIA focussed - but also broader. The Destatis-Hbits contract was a catalyst!
- Lessons learned: explicit need to relate technical and conceptual aspects with (national?) legal requirements. Divergent national viewpoints might be an issue, so lets coordinate national communications.

Original set up

- Initially two phases: qualitative (phase 2) and quantitative testing (phase3)
- Phase 2:
  - Functional tests: functioning of the app
  - Usability tests: user experience
- Phase 2: mostly internal test persons – phase 3: real respondents
- Conditional clause in project protocol: countries can unconditionally opt out of phase 3
- Subthemes: labour time/transport/media use:
  - Interesting use cases (e.g. transport ~ Green New Deal)
  - They mainly serve as POC: can we use this kind of methodology outside of the TUS/HBS ecosystem

Contract

- Original set up was too ambitious – result: split up between phase 2 (design/build phase - qualitative testing - contract 1) and phase 3 (collect phase - quantitative testing – contract 2)

- ~ implicit conditionality: if contract 2 with Destatis –> implement measures already in execution of contract 1
  - prevent gradually stretching up of requirements? How to sell the package?
  - making the case for 'data protection by design'?
  - However: requirement were very general => mismatch between technology and requirements!

- Consequences of the split up:
  - Functional and usability testing (phase 2) can take off
  - However, only testing with internal test persons is allowed!

New set up

- See email communication

- Start functional tests - Bilateral collaboration:
  - Organize log-ins / start testing
  - Bilateral meetings in line with countries agendas
  - Focus group like meetings
  - 4 questionnaires: TUS, labour time, transport, media use

- National legal issues: in principle none left

- Bilateral meetings: prepare next steps:
  - preparation of usabilities test
  - redesigning of the scope

Subthemes

- POC of thematic modularity of MOTUS
- Choice for TUS-related subjects - but with importance on their own (e.g. passenger mobility)
- Four stand-alone questionaires:
  - Drafts – not ready-to-use surveys – ~quickly produced
  - Some innovative elements:
    - Design features : mix of household and individual survey
    - Preload answers in questions: use input from one question in another
  - Testing will be laborious – focus should be on functional elements

WP2.3 HEALTH
the use of accelerometers to measure physical activity

DESTATIS Statistisches Bundesamt · INE Instituto Nacional de Estadística · Insee · Istat · Office for National Statistics · SCB · Statistisk sentralbyrå Statistics Norway · STATEC · Statistics Poland

ESSNET SMART SURVEYS

STATBEL België in cijfers

Research questions

1. What devices are easy to install and use by respondents:
   - compliance with instructions, tolerating the device, tolerating the invasion of privacy.

2. What is the quality of the data:
   - Uninterrupted measurement
   - On intended body part,
   - In the designated direction,
   - During the designated amount of time,
   - Unhindered access to the collected data
   - Respondent evaluation: adherence to protocol, reasons not to adhere, intrusiveness, attractiveness and usefulness of feedback.

3. Who participates, what are considerations (not) to participate, how can we influence that decision

Phase 0: assessment ≈ ✓

- What accelerometers should we use
- On which part of the body should the accelerometer be placed?
- Is one accelerometer sufficient, or should they be appended with e.g., heart rate measurement?
- How do we get data from the measurement device to the backend server?
- How do we make data available for researchers while preserving respondent privacy?

The accelerometer of choice is the thigh worn ActivPal
Heart rate measurement may initially be added, but the goal is to develop algorithms that can do without.

Phase 1: design and preparation of tests ≈ ✓

Tests NL:

1. Feasibility and quality of using people's own accelerometer versus 'our' meter in nonprobability panel n = 50
   - Conditions to participate
   - Sample distribution (demographics and physical activity profile)

2. Feasibility and quality of using people's own accelerometer versus 'our' meter versus cheap incentives offered in probability panel n = 500

3. Respondent willingness, data quality and sample distribution if participants are recruited in a health related Stat Netherlands survey, involving interviewers.

Tests of PL and BE contain elements of the NL tests, possibly with a convenience sample.

Parallel research tract in NL on the topic of accelerometers, with a PhD, involvement of universities (Utrecht, the Hague and Amsterdam), the National Institute for Public Health and Environment, Municipal Health Services and interns, that feeds into the ESSnet.

Statistics Poland

Phase 2:

Test 1 (NL) has been performed, a preliminary paper is available. The accelerometers are now in Poland for their first test. BE will start in 2021, but had not decided when yet.

Phase 2'

- What machine learning models are used in the literature for the type of sensors?
- How to derive and summarize statistics from the sensor data?
- Are there country-specific requirements for infrastructure and logistics?

Deliverables and dissemination

1. report on the surveys and survey questions that can be supplemented or possibly replaced with these sensor measurements.
2. report on respondent engagement and involvement
3. report on shareability across country-specific settings
4. report on infrastructure and logistics specifications as input for WP3
5. open source code, if applicable, and manual for tools deployed in pilots

WP2.4 LIVING CONDITIONS
measuring indoor environment quality

Research questions

- What devices are easy to install and use by respondents, without the need of substantial support and without the need to mount the devices in any complicated way?
- What devices offer the possibility of secure data access
- What devices make it possible to install the device on multiple computers / cell phones (in multiple countries).
- Will it be necessary to (have) build own devices to meet our requirements?
- Do people accept state provided sensors in the private environment? Will they accept using sensors for the purpose of statistics and not for private reasons?

Phase 0: assessment ✓

In this phase, the following activities are performed:

- Literature search into measurement of indoor environment quality,
- Literature search into availability of measurement systems, their accuracy and suitability to yield raw data that can be accessed by the NSI.
- Discussions with stakeholders, universities and other research institutes, like RIVM and TNO in the Netherlands, and VITO in Belgium.

A paper is being written on the results of this phase.

Phase 0 has led to acquiring 25 indoor quality measurement systems: the uHoo air, that measures:

Carbon dioxide CO2, Carbon monoxide CO, Particulate matter PM2.5, Volatile substances VOC, Nitrogen dioxide NO2, Ozone 03, Air pressure, Temperature, Relative humidity

Phase 1: design and preparation of tests

Goal of the first and second test:

1. to get a general idea about the feasibility of doing sensor measurements in people's homes, both from the respondents' point of view, and the NSI's point of view: burden, quality, accessibility, privacy. Respondents from convenience sample or respondent panel (NL-BE-GE)
2. to test the insights from the first test in a probability sample, to additionally test respondent willingness.

Phase 2: tests

NL will perform two tests with 25 respondents each. Respondents place the device in differing parts of their homes. An evaluation phase will be part of the tests. First test will take place in Q4 2020.

GE will perform one extensive test with respondents from test database.

- Respondents are given the device to put in their home.
- Questionnaire on household composition and subjective perception of the indoor environment quality.
- Cognitive interview afterwards about the motivation to participate in the project, concerns, the experience with the device, problems when installing, data safety, possible concerns about sensor data measurement for official statistics

BE will perform one test.

Phase 2'

What machine learning models are used in the literature for the type of sensors?

How to derive and summarize statistics from the sensor data?

Are there country-specific requirements for infrastructure and logistics?

Ethical considerations are of paramount importance in this project

We may measure dangerous indoor quality situations in people's houses.

Should we make them aware? When? After the measurement period or during measurement? Are we liable if we measure but do not warn?

The device has an app that warns if norms are exceeded. Knowledge of the measurements will lead to change in behavior. Do we want that?

Deliverables

- A report on the surveys and survey questions that will be supplemented or possibly replaced with these sensor measurements.
- One report on respondent willingness to participated and reasons for concern.
- One report on shareability across country-specific settings
- One report on infrastructure and logistics specifications as input for WP3
- Open source code, if applicable, and manual for tools deployed in pilots

TIME USE SUBTHEMES

LEGAL-ETHICAL-POLICY REQUIREMENTS

Legal-ethical-policy requirements

- Explicit part of deliverable 2.7 reporting to WP3
- Has been a prominent discussion topic in the ESSnet, especially for WP2.2
- Countries differ in interpretation of GDRP and additional requirements/documentation
- Especially important when data is collected centrally

Legal-ethical-policy requirements

What is new?

- Data collection may use computing and storage options of personal devices
- Part of the data collection may be passive
- Part of the data collection processing may be shifted to the personal devices rather than in-house at the institute
- Part of the collected data may be of specialist content of which part of the respondents has no knowledge, even under informed consent
- In order to be able to participate, one may have to possess a specific device or to accept that one needs to use a device that is provided

Legal documents

In-house data collection:

- Data Protection Impact Assessment (DPIA) extended to cover new data collection features, in particular use of personal (smart) devices for storage and processing

Outsourced data collection:

- DPIA for data collector
- Data processing agreement (DPA)
- Country-specific additional documents: In Germany, for example, the Grundschutz Compendium from the Cyber Security Authority

In WP2.1: Plan is to move from outsourced to in-house data collection
In WP2.2: Hybrid form between outsourced and in-house data collection

Data Protection Impact Assessment

A DPIA is required whenever processing is likely to result in a high risk to the rights and freedoms of individuals. A DPIA is required at least in the following cases: a systematic and extensive evaluation of the personal aspects of an individual, including profiling, processing of sensitive data on a large scale

The European Data Protection Board (EDPB) set criteria for acceptable risk assessments for a DPIA

Legal documents

Important questions:

- What are additional requirements/documentation in the ESS NSI's?
- Do countries put different emphasis/nuance on GDRP operationalization?
- What are the boundaries between legal en ethical requirements?
- What are consequences for an ESS platform?
  - Under in-house data collection
  - Under external data collection

PROPOSAL: The ESSnet NSI's ask their legal departments for necessary additional documentation given the DE example.

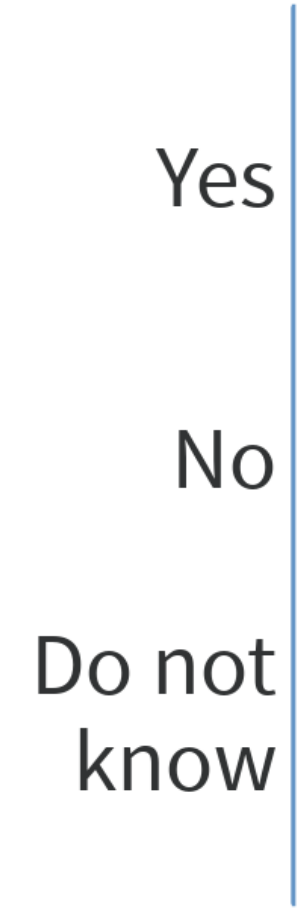
Ethical constraints

Poll through PollEverywhere plugin.

Go to: PollEv.com/barryschoute174

Six questions will be posed on ethical issues for smart features.
Please answer them from the point of view of your NSI.

Do you have an ethical committee for data collection?



Will use of local storage of data on personal devices be acceptable (after consent)?

Yes

No

It depends on the type and amount of data

Do not know

Will local processing of data on personal devices be acceptable?

Yes, after explicit consent

Yes, no explicit consent needed

No

It depends on type and amount of processing

Do not know

To what extent is active involvement required in sensor data collection?

All collected data should be checked by respondents

All collected data must at least be open for respondent assessment

It depends on the type of data

Do not know

How should we involve sample persons when sensor data have specialist content?

Is it acceptable that some persons do not have access to smart devices?

Yes, if coverage rate of devices >90%

Yes, if coverage rate of devcies > 50%

Yes, always

No

Do not know

NEXT STEPS WP2.1 AND WP2.2

Next steps in general

- Start completing D2.7 templates/questionnaires (phase 2/2')
- Collect information needed for phase 3 opt in decision
- Decide whether phase 3 design needs revision
- Coordinate phase 3 design with new HBS-TUS innovation projects

# NSI's	Consumption		Time use	
	NSI's that opted in at start	Other NSI's	NSI's that opted in at start	Other NSI's
5	800	800	2400	2400
4	800-1000	800-1000	2400-3000	2400-3000
3	800-1333	800-1333	2400-4000	2400-4000
6	800	533	2400	1600
7	800	400	2400	1200

NEXT STEPS WP2.2

Next steps WP2.2 – phase 2

- Approach = focus on bilateral and group approach
  - Agile-iterative
  - Cooperative
  - Supportive

- Focus on iterative improvements on different levels:
  - Functional
  - Reusability/modularity
  - Content
  - Privacy by design

Next steps WP2.2 – phase 2

- Functional
  - TUS eco system: hh Q – ind Q – TD – end Q + evaluation Q
  - User experience/Accessibility
  - Iterative => modifications and/or requirement list

- Reusability
  - 3 sub themes
  - Questionnaires + diary refinement

- Inclusion smart data
  - Geolocation data
  - Shaping an inclusive environment/strategy for smart data

Next steps WP2.2 – phase 2

- Content
  - Online questionnaire
  - Online time diary: activity list + methodological settings
  - Contextual questions
  - Translations
  - Metadata

- Privacy by design
  - Front-end: 1 to n application – Client API
  - Back-end: Security/technical/privacy preservation
  - UI/UX: Respondent in the center = interactive approach

Next steps WP2.2 – phase 2

- Step-by-step approach

- ≠ development
  - Using the modular capacity of MOTUS (GSBPM)
  - = inclusive to new developments realised in other projects
    - Source TM
    - Geolocation plugin
    - CRŒSS project (if granted)
    - Own VUB-hbits funding
    - …
  - = inventory of requirements

NEXT STEPS WP2.1

Next steps WP2.1

- Phase 2/2’ evaluation and collection of info for WP3 reporting
- Extension of Stat NL DPIA for storing and processing of data on personal (smart) devices, including English translation
- Opt-in decision phase 3
- Data processing agreements with participating countries
- Phase 3 design

- Tentative: Coordination of phase 3 design with @HBS2

Next steps WP2.1 – phase 2/2’

- An evaluation questionnaire was distributed in July for phase 2/2’
- An extra questionnaire will be prepared for deliverable 2.7 reporting levels
- To be completed in Q4 2020

Next steps WP2.1 – phase 3 opt in

- Deadline is set at Dec 1
- Phase 3
  - Opt-in: LU, NL
  - Opt-out: DE, NO
  - Uncertain: BE, ES, PL
- What info is needed for phase 3 opt-in decision?
- What requirements for phase 3 opt-in decision?

Next steps WP2.1 – phase 3

- Proposal: Stick to original design
  - Experimental conditions: respondent statistics, F2F recruitment
  - Sample sizes:

# NSI's	Consumption	
	NSI's that opted in at start	Other NSI's
5	800	800
4	800-1000	800-1000
3	800-1333	800-1333

- @HBS2 extensions: multi-device, integrated questionnaires, improved receipt scan processing