



# Smart Survey Implementation

## Grant Agreement Number: 101119594 (2023-NL-SSI)

## Work package 4 "Logistics"

***Adapting your business process for smart solutions***

## Deliverable 4.3: Smart advanced stage

**Version 1.0, 2025-04-25**

**Prepared by:**

Marc Houben (Statistics Netherlands, Netherlands)
Marcel Oberije (Statistics Netherlands, Netherlands)
Esther Schepers (Statistics Netherlands, Netherlands)
Joeri Minnen (hbits, Belgium)
Barry Schouten (Statistics Netherlands, Netherlands)
Tim de Jong (Statistics Netherlands, Netherlands)
Fong Yee Wong (Statistics Netherlands, Netherlands)

Work package Leader:
Marc Houben (Statistics Netherlands, Netherlands)
e-mail address : mhf.houben@cbs.nl

# Index

# Executive summary

This is a brief overview of the entire report.

The document is primarily intended for NSIs who are planning to use smart solutions in one or more of their surveys. The main goal of this report is to offer practical tools from a business process perspective. The topics covered in this report are:

- a **classification** for smart solutions,
- a **maturity model** for smart surveys,
- a set of **process building blocks** to be used in the business process,
- a description of **PDCA-cycles** for smart surveys

**Definitions and Classifications**

First, terms like smart data, smart solution, smart feature and smart task are defined. Furthermore, a **classification for smart solutions** is proposed and motivated.

**Maturity model and business functions**

The **goal** of the Smart Survey Maturity Model is

- to help an NSI to assess how mature it is concerning the implementation of smart surveys,
- and to give insight how to become more mature.

The model is useful for all NSI's that are already using smart surveys or are thinking of 'going smart'.

The model consists of five maturity levels and five focus aspects (organization, methodology, business process, IT and logistics). For each cell in the matrix, maturity criteria are described.

This report describes also the relationship between the different maturity levels and **business functions**, also called capabilities. A business function focuses on what the organization must be able to do, to create its products and services. This information is useful for any NSI that is involved in applying Smart Surveys in the organization and wants to gain insight into which business functions are primarily involved in each maturity level.

The model was qualitatively tested within Statistics Netherland and within Statistics Norway. The model can be applied to a specific statistic or to the organization as a whole, depending on the purpose and role of the user in the NSI.

⇨ We state that an NSI is mature enough to use smart solutions in production when (at least) maturity level 3 is reached.

At maturity level 3 one of the criteria is that the business process for production is designed, described and implemented and that the smart aspects are well thought out. The **process building blocks** (as described in this report) help with this.

**Process building blocks**

The main objective of WP4 is to deliver guidelines that will help NSIs to extend their business process to adopt smart solutions in their surveys. Part of this is to help NSIs to model their business process and to identify the capabilities needed. Each NSI has its own situation, its own context, and

its own ideas of applying smart solutions, and thus has its own process requirements. It is not feasible to make a process model for each of these situations.

Building blocks have been devised to help the NSIs provide direction on how they can organize their business process to integrate smart surveys into their surveys. A building block should be seen as a business process activity. The focus is on the statistical production process.

We (only) looked at building blocks that are specifically relevant for using smart solutions.

These can be of two types:

- it's a process activity that is new, so **non-existing for non-smart surveys**.

  E.g. 'Develop app'

- it's a process activity that is **not specific as such for a smart** survey but needs extra attention.

  E.g. 'Train helpdesk employees'

Some process activities are generally usable/applicable, thus **not specific to the type of smart solution.** Other activities are **specific to the type of solution (**E.g. "OCR of the receipt"). In this respect we restricted ourselves to the three SSI use cases HBS, TUS and energy usage meters.

An NSI can use these building blocks to model their production process. Each NSI can vary in the order of the process activities and determine which blocks appear or do not appear in the process. It can be used to model a generic process or a survey-specific process. That is up to each NSI.

And if no actual process model is created, these building blocks can still be useful. It also forms a list of aspects – like a checklist – that you need to take into account if you plan to use smart solutions.

Microservices (see WP3) – as far as in the scope of SSI – have been related to the building blocks. In this way, an NSI gets a clear picture of where microservices can be used in the business process.

Furthermore, for each building block, related actors are identified, i.e., for each process activity, it is made clear what business actor is involved.

**PDCA-cycles for smart surveys**

National Statistical Institutes (NSIs) use PDCA (Plan-Do-Check-Act) cycles to develop and manage smart surveys that adapt to changing technological, regulatory, and user demands. These surveys integrate aspects like IT architecture, methodology, logistics, and legal-ethical requirements to ensure quality and compliance.

Smart solutions are developed through iterative PDCA processes, aligned with frameworks like GAMSO and GSBPM. Each component—such as data collection methods or machine learning models—undergoes its own PDCA cycle to refine assumptions and improve outcomes.

In logistics, PDCA helps design and test new collection strategies, including device-based data gathering. Pilots are used to gather feedback, and results are evaluated against key performance indicators. Lessons learned inform refinements and guide future survey development.

This approach ensures that NSIs maintain high-quality standards, stay responsive to stakeholders, and continuously improve their statistical operations.

**A PDCA-cycle for machine-learning models**

This PDCA-cycle has extra attention in this document.

Smart Surveys aim to reduce response burden by integrating sensor data or assisting the user in completing the survey smartly. In both scenarios, machine learning models are often used to process the data coming from sensors and predict aspects of the respondent's behaviour. In this sense, machine learning models can improve survey efficiency, lessen response burden, or help improve the quality of the response.

Some examples of tasks that can be performed with a machine learning model are OCR, classifying receipt texts or labelling activities. Before a machine-learning model can perform tasks like these, a machine-learning model has to learn to perform these tasks. We call this training a machine learning model.

The data a machine learning model sees in a real-world scenario often changes (and continues to change) from the data it saw during training. In addition, not all data may be available at the time of model development and training. These challenges mean that training a model once will rarely be enough. That is why the machine learning models that we use in smart research will have to be continuously monitored and updated. In the chapter "PDCA-cycle for machine-learning models in smart surveys", we outline a PDCA cycle that will help keep machine learning models up to date. It provides insight into what an organization (NSI) needs to do to maintain machine learning models used in smart solutions.

# 1. Introduction

## 1.1 Objectives of this report

This document provides the end results of the work done in work package 4 "Logistics" of the Smart Survey Implementation (SSI) project. This is the WP4 deliverable of the smart advanced stage, the final stage of the project.

The main goal of WP4 is to deliver concrete **guidelines** that will help NSI's to extend their **business processes** to adopt smart solutions in their surveys. Results should be helpful not only for the parties participating in the consortium for the SSI project, but for all NSI's. The document is therefore intended for Eurostat, for NSIs participating in the SSI project, and for NSIs in general.

The topics covered in this report are:

- **Definitions** of terms (chapter 2)
- a **classification** for smart solutions and smart surveys (chapter 2),
- a **maturity model** with maturity criteria (chapter 3),
- business functions (chapter 4)
- **process building blocks** (chapter 5)**,**
- **PDCA-cycles** (chapter 6).

There are different types of smart solutions (with smart features) that can be applied in a smart survey. An app that has a feature for scanning and processing receipts is an example of a specific smart solution. But use of an energy dongle is e.g. a complete other smart solution. The classification in this report describes a classification of smart solutions. Each type of smart solution has an impact on different aspects[1]:

- Logistics: Different types of smart features need different logistical processes.
- Legal: The different types of smart features have different implications on privacy and security. This is also why in work package 5 'Legal' the taxonomy is used as a framework[2].
- Methodology: Depending on the type of the smart feature, different topics need to be addressed, e.g. the recruitment strategy.
- IT: Different types of IT-solutions are necessary for different types of smart features.

Not all conceivable types of smart features are in the scope of the SSI project. Only three examples are in the scope: scanning receipts using an app, geo-tracking using an app, and using an energy dongle for measuring energy usage. The classification (as described in chapter 2) however describes a broad classification.

The process building blocks are primarily relevant to staff involved in translating a survey design into business processes[3]. Introducing a smart solution into a survey design, will have implications on the logistical activities in an NSI. There will be some new process activities (that are not needed when not using a smart solution) that the NSI needs to implement in the process, and there will be process activities that are not specific for smart solutions but do need extra attention. It also addresses the

[1] These are also the aspects that are addressed in the SSI project

[2] Please see deliverable 5.3 of the SSI project.

[3] Or the other way around: given the business process model, what boundary conditions are there for the survey design?

question which new capabilities are needed when smart solutions are going to be used (e.g. app developing).

This is certainly also relevant for management to get a picture of what is needed in the organisation to start using a smart solution or to grow to a higher maturity level.

For the latter, the maturity model provides support. The model supports management – or the NSI in general – in determining the current maturity level and it provides guidance in order to grow to a higher level. For instance, what is needed to go from 'innovation/piloting' to a smart survey 'in production'.

For this, the maturity model is also mapped to business functions. This guides an NSI in determining which business functions are affected when introducing/using smart solutions. E.g. 'legal, risk and compliancy management' or 'human resource management'. This is particularly important for management and business process engineers of the NSI.

The PDCA-cycle concerning machine-learning models show that maintaining these models, needs continuous investment. It takes effort to keep the models efficient and effective. The organization (NSI) needs to be aware of this.

Bottom line is that this all guides an NSI to determine the impact on the organization from a business processes perspective to adopt smart solutions in their surveys.

## 1.2 The different topics and Reader's guide

In chapter 2 first terms like 'smart solution' and 'smart task' are defined. Second a **classification of smart solutions and smart surveys** is described.

Chapter 3 describes a **maturity model for smart surveys**. The model consists of five levels and five focus aspects. For each combination maturity criteria are listed. In appendix A there is a manual provided that describes how the assessment process looks like and how to use the maturity model.

In chapter 4 we related the maturity model to **business functions**. Each level of the maturity model has different business functions that are primarily involved. This chapter describes the relationship between business functions and the maturity levels.

Chapter 5 describes **process building blocks**. With the use of these building blocks, an NSI should be able to design their own statistical business process in which a smart solution is applied.

In chapter 6, there is a description of **PDCA-cycles**: one looking from the logistical perspective and one looking from the methodology aspect concerning ML-models. A PDCA from the legal-ethical perspective is provided by WP5 (see reference [7]).

## 1.3 References


[1] Smart Survey Implementation; "Workpackage 4: Logistics; Deliverable 4.1: review stage"; Version 1.1, 2023-11-24; https://cros.ec.europa.eu/book-page/report-logistics-41

[2] Smart Survey Implementation; "Workpackage 2: Research Methodology; Deliverable 2.1: Review stage"; Version 1.3, 2023-10-30; https://cros.ec.europa.eu/book-page/report-methodology-21

[3] ESSNet Trusted Smart Surveys (TSSu); "Development of a conceptual framework, reference architecture and technical specifications for the European platform for Trusted Smart Surveys; Deliverable 3.3 Report on the Enhanced Framework"; Version 1.0, 30-03-2022

[4] Smart Survey Implementation; “Workpackage 3: Developing smart data services; Deliverable 3.3: Smart advanced stage report”; Version 1.0

[5] ESSNet Big data II, Deliverable F1: BREAL - Big Data REference Architecture and Layers - Business Layers, 9 December 2020, Link

[6] GSBPM 5.1: https://statswiki.unece.org/display/GSBPM/Clickable+GSBPM+v5.1

[7] Smart Survey Implementation; “Workpackage 5: Design level Legal-Ethical; Deliverable 3.3: Smart advanced stage report”; Version 1.0

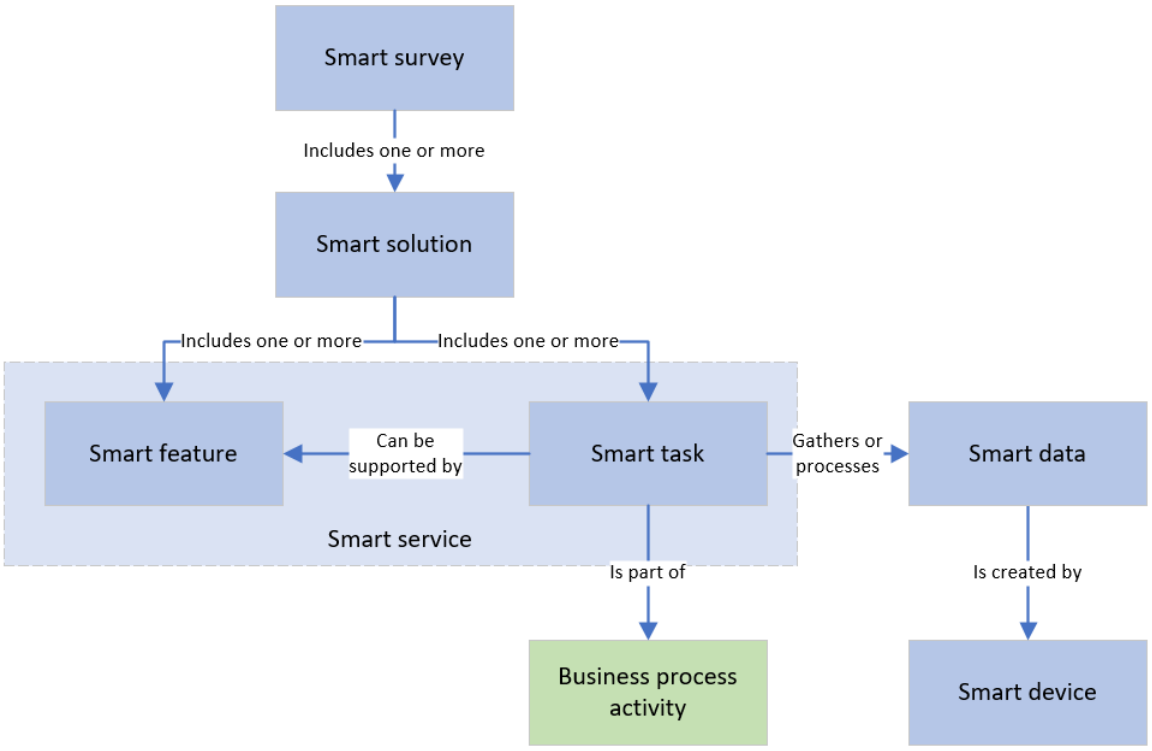
# 2. Definitions and Classification of smart solutions

## 2.1 Introduction

The need for clear terminology and in particular a classification of smart solutions has two motives. One is to be able to generalize to other applications of smart surveys not in the scope of SSI. Another is to facilitate communication between different design and analysis disciplines underlying smart surveys. The goal of a classification is to create categories of smart solutions that are sufficiently homogeneous to develop generic methodology[4], logistics[5], IT[6] and legal-ethical procedures[7].

Before defining a classification of smart solutions, we need to define what the term smart solution means and how it relates to other terms like smart features and smart surveys.

## 2.2 Definitions




**Smart data** is data that is created by a smart device. If smart tasks are applied to this data, it is still called smart data.

A **smart device** is an electronic device that is equipped with advanced technology that allows it to collect, process, and share data, often enabling remote control, automation, and interaction with other devices or systems. The device has therefore the capability to connect to the internet and to

[4] e.g. push-to-smart recruitment and motivation strategies, active-passive trade-offs in respondent interaction, machine learning predictions and other data science methods
[5] e.g. monitoring and analysis, interviewer training, (re)training of classification models
[6] e.g. clustering of tasks within services, input-output specifications
[7] e.g. privacy risks, mitigation measures

other (smart) devices. The device has capabilities to collect data from its usage (e.g. take photo of receipt) or its environment (e.g. temperature of the surrounding), in principle by the use of sensors.

A **smart task** is an action to *gather*[8] smart data or to *process*[9] smart data. A task can be fully automated, semi-automated or manual. The actor can be someone in the business but can also be the respondent. A smart task is part of a business process activity[10]. In general, a process activity exists of one or more tasks.

A **smart feature** is a functionality of an IT-system or smart device that supports a smart task.

A **smart solution** is the totality of smart tasks (with the support of smart features) needed to gather and process smart data to contribute to the goal of the survey.

A **smart service** is an IT-component that comprises one or more smart features (and thus tasks). It is a kind of package that can be used by an IT-system to perform its task.

A **smart survey** is a survey that uses one or more smart solutions.

To fix thoughts, three examples of smart solutions:

- Receipt scanning for household budget surveys: The smart device is the smart phone of the respondent. The camera is the smart feature. The smart data are pictures of receipts. Smart tasks are, among others, respondent checks on the picture, image pre-processing, text extraction, text interpretation, classification.
- Location tracking for time use surveys: The smart device is the smart phone of the respondent. The GPS-sensor is the smart feature. Smart data are location data. Smart tasks are, among others, checking and supplementing location data, filtering/smoothing, stop-track segmentation, predictions of purpose and transport mode.
- Energy data donation through a dongle: The smart device is the smart meter itself and the dongle that is connected to the smart meter. The smart feature is the capability of the dongle to collect the energy data from the meter and to communicate with other devices. Smart data are energy usage and/or energy production data. Smart tasks are, among others, checking and supplementing data, prediction of type of energy consuming devices being used, classification of devices.

## 2.3 Smart tasks

The core of smart solutions, thus, consists of smart tasks.

The following properties can be given to a smart task:

- TYPE: Collecting, cleaning, enriching, editing, imputation, prediction, transformation
- TIMING: In-survey, post-survey

[8] You could also call this *to collect* or *to create*

[9] The term processing covers e.g. also the storage of the smart data, connecting to a source with smart data, and interpreting or contextualizing smart data.

[10] Note that also when the task is performed by the respondent, we call it a business process activity.

- SOURCES: No additional sources, in-house admin data, public online data, ACTORS: Fully automated, semi-automated with respondents, semi-automated with staff, semi-automated with respondents and staff
- UI: No respondent interaction needed, in-platform respondent interaction, in-service respondent interaction
- RULES: Set of parameters that are input to the task
- STANDARDS: Lower pre-specified thresholds to quality of smart data
- TOOLS: Stand-alone, dependencies on existing software/tools
- LEARNING: Output contains no data for learning, output contains data to learn per respondent, output contains data to learn across individuals, output contains data to learn per respondent and across respondents

The list of properties is not necessarily complete.

A brief explanation of the properties:

*Type* of task refers to the operation performed. Being processing steps, smart tasks are one of the following:

- Collecting: The actual collecting of the smart data
- Cleaning: Smart data noise and outliers are removed
- Enriching: External data are added
- Editing: Smart data are confronted with edit/plausibility rules and edited if needed
- Imputation: Missing smart data are imputed
- Prediction: Smart data are classified into a specified categorization
- Transformation: Smart data are aggregated, combined and/or fused

*Timing* describes the choice between doing a task during or after data collection. There are two motives for doing a task in-survey, i.e. during data collection: respondent interaction to improve quality and respondent feedback to improve engagement. However, an NSI may decide to implement an in-survey task as post-survey, e.g., abstain from feedback or interaction.

*Sources* are additional data about the respondent or groups of respondents and are used as features in methods to clean, edit, impute, predict or transform. They come in two forms. The first is through data already in possession of the institute. In SSI context, this will often be administrative data. The other option is linkage of public online data.

*Actors* conform to human-in-the-loop which can be respondents themselves and/or data collection staff.

When a human-in-the-loop component is included, then the *UI* must be set at either in-platform (i.e. outside the service itself) or in-service (i.e. the service has a UI itself). The difference is that in-platform is a solution-specific responsibility and decision, e.g. the MOTUS and @HBS app may handle these differently.

*Rules* depend on the type of task. They are input parameters to methods employed within the task.

*Standards* are lower thresholds to smart data quality. Below these thresholds the task is/cannot be performed. They imply that a preceding smart task is re-initiated, or that respondent context or supplements are imperative.

*Tools* are also dependent on the type of task and are parts of the methods that are external. External means that external libraries/packages are included, described and validated in the literature.

*Learning* refers to the decision to use output data of the smart task to adapt the applied AI-ML to individuals and/or to groups of respondents. The form depends again on the type of task. If true, then the smart task output (and thus also the services including the smart task) need to specify what part of data needs to be separated for learning purposes. Learning can be active but also online.

While general properties can be identified for smart tasks, their format and implementation depend on the type of smart feature(s). It is, therefore, useful to study commonalities in smart features.

## 2.4 Accuracy gap and output gap

The most influential finding in SSI is the importance of the so-called accuracy gap and output gap of smart data. The accuracy gap is the distance between ideal smart data and smart data as obtained in practice.

- ⇨ An *accuracy gap* means that smart data need to be adjusted, i.e. one or more smart tasks are imperative.

It is the counterpart to the output gap which is the distance between desired data (based on output specifications) and the ideal smart data.

- ⇨ An *output gap* implies that smart data need to be supplemented and/or adjusted.

An example to fix thoughts: Location data stemming from tracking contain gaps for a variety of causes (tall buildings, underground, lost signal, device in hibernation, battery depleted), outliers (switches between GSM-GPS-Wifi, low accuracy of GSM, implicit tweaking of data by native app code) and random noise (sensor imprecision). All these add up to an *accuracy gap* between perfect location tracking and actual location data. The *output gap* goes two ways: There is a surplus of information. Location data, when enriched with contextual data such as OpenStreetMaps, gives more detail than needed. Once predictions of travel mode and travel purpose have been made, the exact routes travelled are irrelevant. So are the addresses and names of visited stores. Also, the speed with which was travelled is no longer of of interest. There is also a shortage of information. Even when locations would be perfectly measured, location data are insufficient to derive all activity types; at home several activities take place with no other difference than the time stamp. Also, location data cannot reveal with whom an activity was done, whether there were side activities and how much the respondent enjoyed the activity. While a surplus of information calls for data minimization, a shortage of information implies additional consultation of respondents. The two gaps are connected: The larger the accuracy gap, the stronger the need to employ the full detail of smart data for adjustment and editing.

As output gaps depend fully on the information need of the application, their prevalence and influence cannot be part of the classification of smart solutions. However, while the impact of an accuracy gap also, to some extent, depends on the application, it is a property of a smart feature.

- ⇨ Low accuracy of smart data influences the design of methodology, the design of IT-solutions, the design of the logistical process and the decisions made in legal and ethical boards. For this reason, it must be included in the taxonomy because smart features with small to negligible accuracy gaps need to be handled very differently from those that may have big gaps.

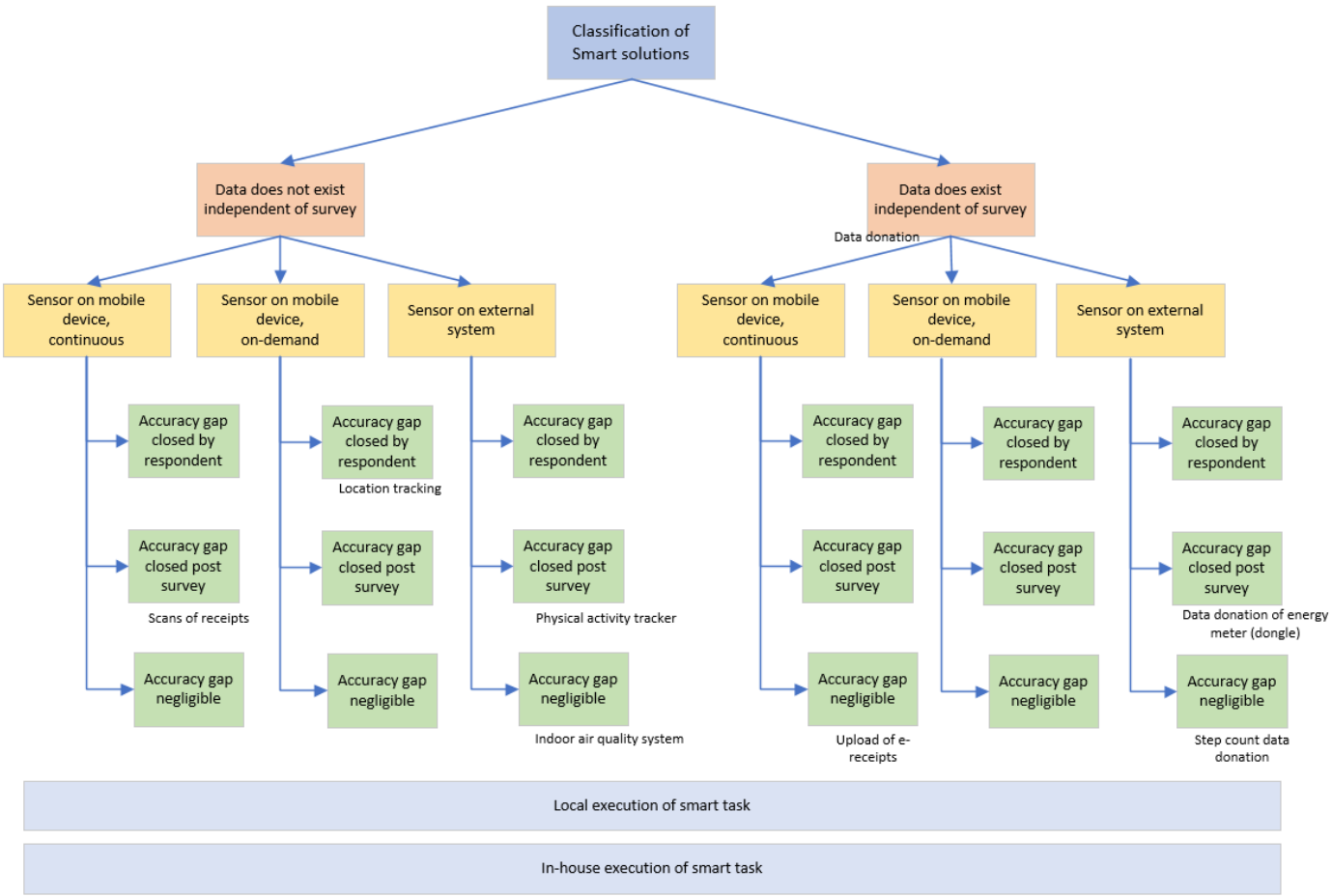
Another important finding is that, once the accuracy gap is evaluated and rated, further specification of the type of sensor does no longer seem to be (very) relevant.

## 2.5 A classification of smart solutions

The objective of the classification is to create classes that are sufficiently homogeneous for all perspectives: methodology, IT, logistics, legal-ethical. The goal of a class is that something general can be said - from each of the perspectives - about the implications of using that type of smart solution. In concrete terms this means that if an NSI is going to use a solution from another class, the NSI needs to look carefully into the consequences and probably needs to design a new solution. On the other hand, if the NSI is going to use a solution that is of the same class of another already used solution, the available solution can probably be used without much adaptation.

Smart solutions of the same class may have relatively similar:

- Respondent workflows
- In-house logistics
- IT architecture
- Back-offices
- AI-ML and UI-UX respondent interaction
- Legal-ethical decision rules




The first distinction we made is whether the **data already exists** prior to (thus independent of) the survey.

- Data does exist: There is an existing data source at the level of individual population units with the data. That means there is a form of data donation in the solution.
- Data does not exist: In this case the data will be gathered specifically for the survey[11].

This choice is based on the answer on the question whether or not there is an existing data source available, that fits the requirements of the survey. This is primarily driven by methodology.

The second distinction determines what **type of measurement** (sensor) is used.

- Sensor on mobile device: This means that the collection of the data is on the device that is handled by the respondent. So, important parts of the handling of the data are done on this same device (e.g. the app runs on this device). E.g. a smart phone.
- Sensor on external system: This means that the collection of the data (use of sensor) is on another device than the handling of this data. So, the sensor is not on the same device that is handled by the respondent. E.g. an activity tracker.

[11] Or maybe a small set of surveys.

With the first option, sensor on mobile device, we make an additional distinction between continuous measurement and measurement on-demand. Continuous measurement means that the sensor is measuring continuously during the reporting period. Consent to employ the feature needs to be asked from respondents. However, once consent is given respondents can retract it but are not asked continuously whether the consent still holds. Examples are location tracking, indoor climate, physical activity tracking and energy usage. The counterpart is on-demand measurement, where the sensor is only used when instigated by the respondent. Examples are receipt scanning and e-receipt uploading. An on-demand measurement implies, by definition, an active respondent involvement. A continuous measurement could be implemented as fully passive, once consent is given. However, in most applications some minimal active involvement is needed.

This gives us six combinations:

| | | |
|---|---|---|
| Data does not exist independent of survey | Sensor on mobile device, on-demand | The data is gathered by a smart feature of a (general) smart device that is handled by the respondent. E.g. a smartphone.<br><br>And where the respondent has to do some action to gather the data. |
| | Sensor on mobile device, continuous | The data is gathered by a smart feature of a (general) smart device that is handled by the respondent. E.g. a smartphone.<br><br>And where the data is automatically and continuously gathered. |
| | Sensor on external system | In this case a smart device/system is used that is not handled by the respondent to gather the data. |
| Data does exist independent of survey | Sensor on mobile device, on-demand | In this case the data would exist prior to any survey, whereas the data is gathered by a smart feature of a smart device handled by the respondent, and where the respondent has to do some action to gather the data.<br><br>That is a form of data donation. |
| | Sensor on mobile device, continuous | In this case the data would exist prior to any survey, whereas the data is gathered by a smart feature of a smart device handled by the respondent, and where the data is collected continuously.<br><br>That is a form of data donation. |
| | Sensor on external system | In this case the data would exist prior to any survey and the device to gather the data would not be handled by the respondent. |

The third distinction is how the **accuracy gap** is closed.

- The accuracy gap is closed with the help of the respondent[12].
- The accuracy gap is closed post survey, by the NSI itself.
- The accuracy gap is negligible, so no specific actions have to be taken to close the gap.

This choice is based on the answer on the question to what extent a gap is acceptable, and if not acceptable, to what extent the respondent can be involved and to what extent the logistical process gets too complex. The choice is primarily driven by methodology and logistics.

The fourth and last distinction is between local and in-house **execution of smart task(s)**. Part of the tasks may be performed on the respondent device. An example is evaluating and adjusting the quality of receipt scans/photos. Another part of the tasks may be done in-house such as classification of products and visited locations. The classification is entirely local, local and in-house, and entirely in-house. The choice where to perform tasks is influential in IT, UI-UX and in legal assessments.

How does this affect the four different perspectives?

Data does or does not exist independent of survey:

- Methodology:
    - Data that exist independently of the survey lead to a data donation process that needs to be explained to respondents and respondents need to be able to control.
    - Recruitment-motivation: In recruitment strategies, the donation of data must be stressed. This is for one because it may reduce burden. However, also there must be clear and transparent information on how data are handled and whether there are alternatives.
- IT:
    - Authorization protocol: Since data were collected for other purposes than the production of survey statistics, they will in general only be accessible through a prescribed procedure. This procedure needs to be implemented in the application.
    - Service: The service performing the smart data collection, and possibly execution of one or more smart tasks, needs to communicate to an external source. This implies it needs to conform to data models of the donated data through an API provided by the data holder. Also, data protection measures may need to be adapted as respondents need to see what is being collected. Consequently, there is data transfer to the respondent's device.
- Logistics:
    - Each option leads to a different logistical process: a different process design is necessary. Also monitoring of the process will have different needs.
- Legal-ethical:
    - Whether the same data were collected by the NSI or by another party could be a topic of discussion for an ethical committee. However, using existing data over new data will always be favourable and, hence, not be influential. Data donation comes with new data security risks, but these are viewed as part of the IT design level.

[12] However, maybe some action to close the gap is additionally done post survey

Sensor on mobile device or on external system:

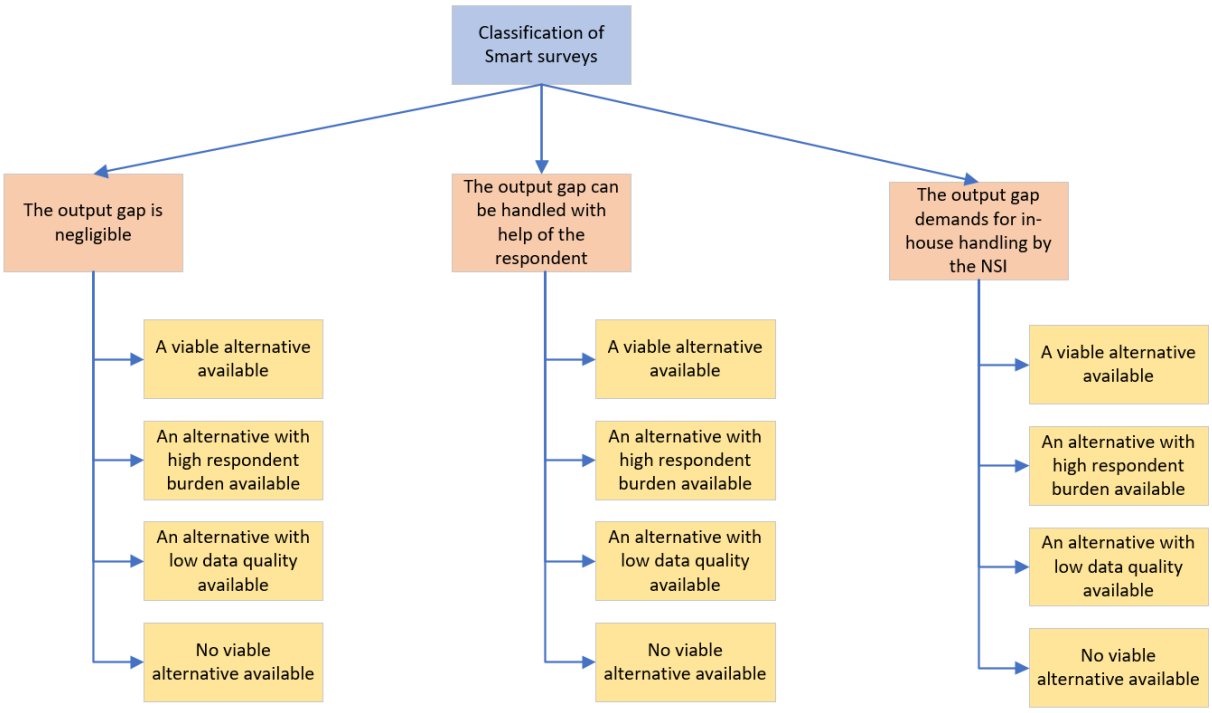
- Methodology:
    - Communication with respondents will depend heavily on the device used. An internal mobile device sensor needs to be initiated, handled and checked by respondents (starting the location tracking, taking a picture). External sensor systems data need to be combined and/or imported into the smart survey back office and if communication is needed into the application. Instructions on how to do this need to be embedded in the UI. Also, the respondent workflow will differ.
    - Recruitment-motivation: The employment of the type of measurement needs to be explained beforehand. It may reduce burden but also bring new burden. It requires clear and transparent instructions and information. Additional burden or complexity to respondents must be avoided as much as is possible.
    - AI-ML: Data science procedures will depend heavily on the type of data.
    - The local handling of smart tasks may affect computational power and computation times, but other than that does not strongly affect methodology.
- IT:
    - Frontend: A sensor measurement has a big impact on how an application operates and how the solution/platform is implemented. Applications need to communicate to sensor systems that often come from a broad variety of brands/models. Maintenance of applications depends heavily on updates and innovations in smart devices and app store requirements.
    - Backend: The (smart) data coming in from the different types of measurements hugely differ in terms of format, time resolution and size.
    - Data security: Local, in-device handling of data requires stronger data security measures.
    - Service: The development and integration of services heavily depends on where smart tasks take place.
- Logistics:
    - The logistical process will look differently. When data is gathered (and/or processed) on an external device, the data should be transported from that device (Bluetooth, USB, API).
    - Monitoring: The type of measurement adds an entirely new dimension to monitoring of incoming data.
    - Helpdesk: As respondents may be new to different types of measurements, the (technical) helpdesk needs clear FAQ and technical solutions.
- Legal-ethical:
    - Ethical evaluation and data minimisation: Ethical committees require evaluation of the burden put on respondents and the added value of new types of data. Furthermore, ethical committees will evaluate the potential impact on the public image of the NSI.
    - Data security: While it is up to IT to secure data collection and handling, risk assessment for data processing outside NSI environment is new and receives special attention in forming DPIA's. In time this dependence may, however, vanish.
    - Privacy-by-design: Local handling of data may be one of the solutions to guarantee data minimization. Only part of the data is really submitted and, therefore, conform more strictly to GDPR.

Accuracy gap:

- Methodology:
  - UI-UX: When respondents need to provide assistance then this can only be organized through the UI.
  - AI-ML: Data science procedures are heavily affected by errors in data but are themselves also the main solution.
- Legal-ethical
  - Privacy-by-design: The larger the accuracy gap, the more pressure on data minimization and privacy-by-design decisions.
  - DPIA: Smart features that lead to inaccurate smart data require experimental data collection in order to learn and let procedure converge. In practice, this means data protection impact assessment are made in two steps. First an analysis DPIA is made that leaves more leeway. Second, a production DPIA is prepared that will stress data minimization as much as is reasonable.
- IT:
  - Backend: When data can be adjusted by respondents, database management and backend operations change drastically. Version management and reproducibility are key and impact the construction of data models and database management.
- Logistics:
  - There could be an impact on human-in-the-loop procedures to maintain AI-ML.

## 2.6 A classification of smart surveys

When adding the context of a specific set of survey output/statistics, two further application-dependent distinctions are made.



The first application-dependent distinction concerns the **output gap** that was introduced in Section 2.3. Three settings are distinguished:

- The output gap is negligible, and no action is needed
- The output gap can be handled with help of the respondent
- The output gap demands for in-house handling by the NSI

The categories reflect the need for smart data to be stored and processed in-house. In the case of a negligible output gap, the submitted data are (almost) exactly what is needed. This setting rarely holds. If not true, then the question follows whether the respondent could resolve the output gap during reporting. This would still pose a favourable setting as any surplus of information would be temporary only. If (supplemental) in-house evaluation is needed, then the surplus of information may have to be accepted.

The second distinction is on the **availability of viable alternatives** to the smart feature(s). This distinction was added in the last SSI project stage based on field test findings. A viable alternative is one that reaches accurate smart data at an acceptable respondent burden. Four categories are distinguished:

- A viable alternative available,
- An alternative with high respondent burden available,
- An alternative with low data quality available,
- No viable alternative available.

The distinction was added as it directly appeals to the logic for the respondent. To fix thoughts: Receipt scanning for HBS has a manual data entry alternative known to be burdensome. Energy data donation has no viable alternative. Location tracking for a travel survey has a manual alternative with low data quality. Location tracking for HETUS has a viable manual alternative.

As for smart features, the impact of the two criteria may be evaluated for the survey design levels.

Output gap:

- IT:
    - Frontend: When smart data need context or supplemental data, then the frontend need to be able to combine, or react to, different types of data.
    - Backend: An output gap leads to a more complex database structure, including a mix of smart data and Q&A data, where the Q&A data definitions may be relative to the smart data.
- Legal-ethical:
    - Ethical committee/review: Ethical decisions strongly depend on the added value and the burden put on respondents. The larger the output gap, the more argumentation is needed to justify a smart feature.
    - DPIA: Output gaps often require storage of smart microdata for further handling, with or without help of a respondent. Acceptance of subsidiarity and proportionality decisions and privacy-by-design in a DPIA strongly depend on such data flows.
- Methodology:
    - UI-UX: The UI needs to combine both smart data and survey data.

- Logistics:
    - No procedures seem to be affected by the prevalence of output gaps.

Alternative(s) to smart feature(s):

- IT:
    - Frontend: The UI should allow for multiple data entry options in a clear and instructive way.
    - Backend: Multiple data entry options lead to a more complex database structure and monitoring of incoming data.
- Legal-ethical:
    - Ethical committee/review: The availability of an alternative has a favourable impact on ethical assessments. The respondent has a choice.
    - DPIA: The availability of an alternative has a favourable impact on risk assessments. Any output gap can be avoided by the respondent. In case no viable alternatives are available, then it will be easier to accept an output gap.
- Methodology:
    - UI-UX: The UI-UX needs to be adjusted in order to allow for different streams of data.
    - AI/ML: AI/ML training and retraining may have to be adapted to any selection in what data is entered through the smart feature.
- Logistics:
    - Monitoring: The monitoring of incoming data may be affected by different data entry options.

# 3. Maturity model

## 3.1 Introduction

The terms maturity framework, maturity model, maturity criteria and their relationships were described in the WP4 deliverable 4.1 (see reference [1]).

Maturity models – in general – are defined in terms of a progression through levels that describe process characteristics. When an organization gains an understanding of process characteristics, it can evaluate its level of maturity and put in place a plan to improve its capabilities. With each new level, process execution becomes more consistent, predictable, and reliable. Progression happens in a set order. No level can be skipped.

The **goal** of the Smart Survey Maturity Model (SSMM) is:

- To help an NSI to assess how mature it is concerning the implementation of smart surveys,
- To give insight how to become more mature.

The Smart Survey Maturity Model can be used to evaluate smart surveys overall within the organization (NSI) or can be used focussing on a single smart survey implementation (for example for HBS).

An NSI can have different business drivers to conduct the maturity assessment:

- **Organizational readiness for process improvement**: In this case an NSI recognizes a need to use smart features in data collection and begins by assessing its current state.
- **Governance**: In this case the NSI needs more guidance in the governance of smart surveys. A maturity assessment can thereby be helpful for planning and establishing governance concerning smart solutions.
- **New technology**: Advancements in technology offers new ways to use smart features. In this case the NSI wants to understand the likelihood of successful adoption.
- **Smart survey issues**: In this case the NSI is confronted with smart survey issues, and the NSI wants to assess its current state in order to make better decisions about how to address and solve the issues.

Earlier versions of the maturity model were evaluated by Statistics Netherlands and Statistics Norway. Their feedback was very valuable to compose the version of the model described in this report.

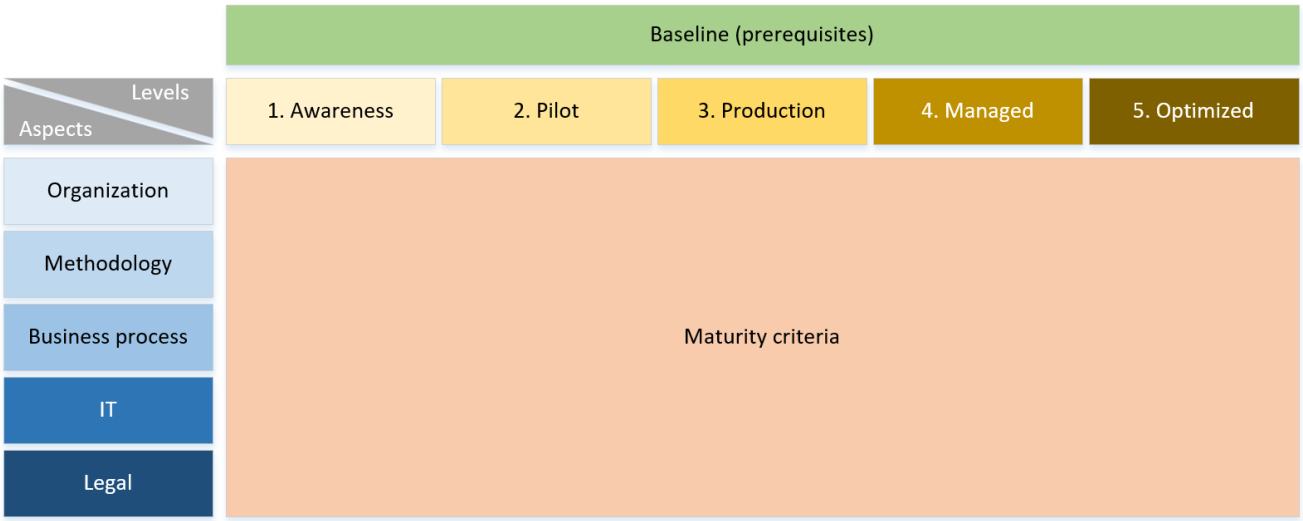
## 3.2 Model

Ideas for the structure of this model were taken from existing maturity models such as Capability Maturity Model Integration (CMMI), Business Maturity Model (BMM), Integral Data Analysis Model (IDAM), Business intelligence (BI maturity model), Service Excellence Maturity Model (SEMM). These models each have their own focus area. We did not find a model with focus on smart surveys. However, we were in need of a model with that focus.

Almost all of these models distinguish 5 levels of maturity. We went with that, just because 4 levels seem to lack sufficient growth possibilities and more than 5 would be too complex for our purpose.

Because our focus area is the statistical business process, we were looking for levels that allow us to describe the innovative process of applying smart solutions. That process starts with generating

ideas, then trying these ideas in pilots and deciding whether or not to apply the smart solution in production. When having a first process/solution in production, you could continue further developing your production process by managing and optimizing your methodology, process and solutions.




*Framework of the smart survey maturity model*

To get started with Smart Surveys, you need to be in a situation where you meet a number of preconditions. We call this the Baseline. In addition, we distinguish the 5 maturity levels in our model: Awareness, Pilot, Production, Managed and Optimized. Each level is described by a general characterization. Furthermore, for each level, maturity criteria have been described. These criteria are described from 5 different aspects: Organisation, Methodology, Business process, IT and Legal.

**The 5 levels, with some terms to characterize it:**

- Awareness

  *Pioneer, thinking, beginning, learning, literature*

- Pilot

  *Exploratory, investigating, starting, start-up, business case*

- Production

  *Implemented, used in production process, applied, separate solutions*

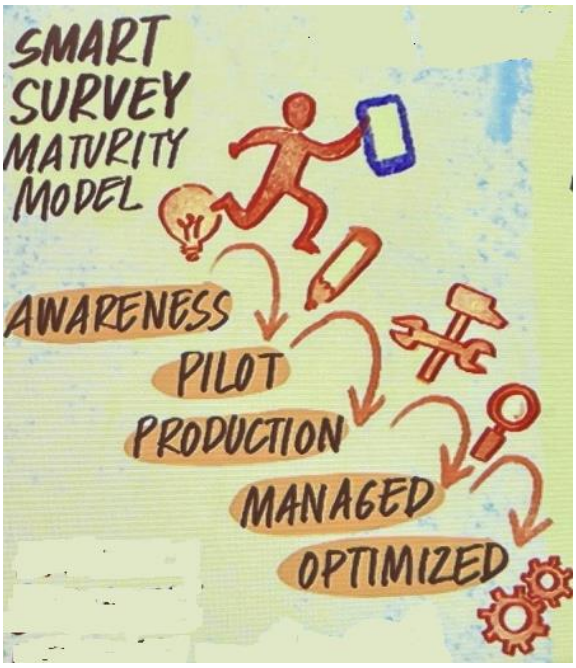
- Managed

  *Standardized, in control, integrated, shareable solutions*

- Optimized
  *Continuous improvement, excellence, best in class, fluent*

Appendix A contains a manual for using the model. Using that, an NSI should be able to conduct an assessment using the maturity model.

## 3.3 Maturity levels with maturity criteria




### 3.3.1 Baseline

The Baseline should not be seen as a level. It describes some **preconditions** that should be met before it is meaningful to use the model and/or to start thinking about 'going smart'. So, check based on the conditions mentioned below whether you 'tick' the boxes. If you don't tick all the boxes this should be seen as a warning that maybe you are not ready to start using smart solutions and using this model will not add much value. However, this is not hard science. The model is not perfect and if you think that there is no obstacle, of course feel free to use the model and assess your maturity.

1. Smart surveys make use of smart devices, which need a kind of network (e.g. GSM-network) to communicate. There should be sufficient coverage of this network in your country. If larger parts of your country lack this coverage, it will restrict the possibilities you will have using smart surveys.
2. Staff: The organization has one or more IT teams in place, responsible for managing and maintaining IT tools and services.
3. Staff: The organization has Methodologists (Research and Development) on staff.
4. Innovation power: The organization aims to innovate in the field of surveys and has some basic innovation power.
5. Innovation power: Resources are available or can be made available for innovation of IT tools.
6. Smart surveys mostly use mobile devices. If your NSI does not yet use a mode where mobile devices can be used, it could be hard to jump directly into smart solutions. So, we would state that you should already be using the CAWI mode and that CAWI questionnaires are usable on mobile devices.

### 3.3.2 Awareness

**General characterization**

At the awareness level, an idea is initially suggested to use smart surveys as an observation mode. It is suspected that smart surveys can help with some issues that an NSI encounters (for example, to achieve better quality responses) or where an NSI sees opportunities to improve the process (for example, a more efficient process). But whether this idea will be successful is unknown. Initially, the focus is on providing a global orientation: what you will encounter with smart surveys. Fundamental choices still have to be made, but it is unknown which decisions will have to be taken. The NSI is just starting out, with little to no experience concerning smart solutions. Within the NSI there is little or no guidance. Much of the activities are done on gut feeling, not yet on well thought out plans.

Maybe an exploratory study is being carried out for the idea to see what is known about the preconditions that apply: what is and what is not allowed/possible.

**Organization**

1. Methodologists (Research and Development) are in the lead. It leans on individuals. Some methodologists are very enthusiastic. Others sceptical. The rest of the organization is rather unaware.
2. The organization is still debating whether smart surveys have a future. There is no business case and not much of a business vision yet.
3. The organization is still thinking in which situations smart surveys would add value.
4. Maybe initial (small) proof-of-concepts are carried out, but not yet pilots with respondents.

**Methodology**

1. Methodology studies literature and available use cases for orientation, to form a first view, to form a first opinion.
2. The methodology to be used is still unclear. First ideas emerge.
3. Some basic standalone proof-of-concepts are devised.

**Business process**

1. There is no business process in place concerning smart surveys. The existing business process is not yet changed.
2. Business process designers are not yet involved.

**IT**

1. IT is focused on proof-of-concepts concerning IT-aspects or on supporting proof-of-concepts of Methodology.
2. IT development is very limited and rudimentary.
3. A rudimentary smart solution(s), like an app, with very limited functionality may be available/developed.
4. Knowledge about e.g. developing an app is limited. There is (hardly) no experience developing an app.

**Legal**

1. Knowledge about the legal, ethical and privacy aspects of smart solutions is non-existing or limited.
2. There is no DPIA for smart solutions.
3. The preconditions from a legal and ethical point-of-view will be explored for the first time.

### 3.3.3 Pilot

**General characterization**
There are uncertainties about Smart Surveys that need to be further investigated first. The idea of using smart surveys is put to practice. The riskiest uncertainties are first identified and eliminated by carrying out proof-of-concepts. More roles are involved at this level, although the number of roles is still limited. While the previous level (awareness) was mainly focused on the idea of smart surveys (not how to apply the solution within the NSI), this level focuses more on the application of the solution within the organization. However, the application is not yet so specific that a statistic is created with the result, but more to gain insight into what needs to be done in order to be able to create statistics later in the next level. There may be informal collaboration, possible due to limited organisational size. The business processes are ad hoc, scattered and unconnected. The NSI is still searching for a business case, searching for 'good' methodology, IT possibilities and legal restrictions. The goal is not to make statistics yet.

**Organization**

1. The organization does see some potential in using smart surveys and wants to investigate it further.
2. The innovation department (R&D) has the lead, drives the initiative. The focus is on methodology and IT, paving the routes for future projects.
3. Smart solution, as a subject, is mentioned in the vision of the NSI or at least in the vision of the responsible department. The vision says that the subject should be investigated.
4. There are sufficient money/resources available to carry out one or more initial pilots to investigate smart surveys. The organization is prepared to make funds available for this.
5. A communication plan is drawn up and implemented. The plan determines which and how to involve stakeholders within and outside the organization.
6. Focus is on the question whether there is a positive business case for using smart surveys.

**Methodology**

1. Alternative strategies are tested in field pilot(s), e.g. recruitment strategy, motivation strategy, etc.
2. It is examined what information can be fed back to a respondent.
3. Comparisons are made between smart and non-smart strategies, e.g. regarding response rate and representativity.
4. Human computer interaction is being investigated. Pilots are used to test UX/UI.
5. In-app respondent interaction is being evaluated and optimized through usability tests.
6. Machine learning routines are being developed and tested. The achievable quality is examined.

7. A goal is to prove that the methodology can eventually be raised to a level so that maturity level 3 can be achieved.

**Business process**

1. Fieldwork is monitored via a minimum set of indicators based on the process data. Process data is collected; however not per se in a standardized way.
2. The process to be piloted is tested and accepted as a chain.
3. The process is ad hoc. And probably largely with procedural/manual activities.
4. Business process designers are slowly involved and asked to determine what consequences there are for the business process when smart solutions need to get to production.

**IT**

1. There is growing knowledge about developing apps/smart solutions.
2. There is a discussion whether apps and (micro)services should be developed by the NSI or by an external party. Is it a capability the NSI wants to get itself?
3. Smart solution(s) with limited functionality are available on main devices and operating systems (like Android and iOS).
4. Some mainstream (personal) devices are available for testing purposes.
5. The testing infrastructure to test apps and smart services is rudimentary.

**Legal**

1. Knowledge about the legal, ethical and privacy aspects of smart solutions is present. Guidelines/recommendations have been gathered from (inter)national authorities.
2. There are no ethical objections about what smart solutions are used for, how they are used and with whom. In case of doubt, the ethics committee is asked for advice.
3. A DPIA (including risk analysis) is available for each smart solution pilot and assessed by a legal officer. The pilot complies with the applicable rules and legislation. Thereby the privacy-by-design choices for the smart solution pilot are motivated, documented and approved.
4. However, there is no DPIA for actually creating statistics with smart solutions.
5. Risks (e.g. security) in pilot(s) with smart solution have been assessed, evaluated and sufficiently mitigated according to the relevant officers (e.g. chief security officer).
6. Informed consent procedures for the smart solution pilot (like the use of sensors) are tested, optimized and compliant with legal requirements.

### 3.3.4 Production

**General characterization**

The organization is confident that it can create statistics with smart survey data, however the focus is probably on only one smart solution (e.g. receipt scanning). The process, IT[13], methodology and legal aspects are not standardized or generic, or in other words, are specific to one statistic. Processes are **stovepipes**, lot is custom-made. However, the whole process (data collection, processing, analysis) is covered so that statistics can be made. All departments that produce statistics with smart survey observations are involved, although the focus of the staff is within their

[13] However some IT-components like microservices are generic

own domain (part of the chain). The organization is in control to the extent that the process documentation has been described and a simple PDCA cycle (see chapter 6) has been set up for one specific statistic, where the improvements are reactive. This means that an incident must first occur before the process will be improved. Improvements are primarily focused on the survey at hand. The PDCA is primarily organised per study, not across studies.

**Organization**

1. The decision has been made to apply a smart solution for at least one survey. So, there is a positive business case. The organisation had determined that smart surveys have added value.
2. The organization has developed new capabilities needed for the smart solution being used.
3. The production department has the lead. The focus is on: “the statistic has to be made using the smart solution”.
4. The organization had/has a ‘change plan’ how to move from an innovation situation to a production situation.
5. Focus is not yet on the full chain of the business process, but on the individual departments.
6. The vision regarding smart surveys is communicated, is known within the organization and is acted upon within the organization.
7. The tasks, responsibilities and authorities are known, concerning the production process. It has been decided where the new tasks and services will be assigned in the organization.
8. Relevant personnel have been trained and has the necessary knowledge for conducting the concerning surveys. However, knowledge is not yet widespread in the organization.
9. Agreements (SLA, products, services catalogue, etc.) are available for the services that can be used.

**Methodology**

1. The methodology that is used was proved to be effective.
2. For the smart solution(s) at least the following has been taken care of:
    a. used methodology is recorded and documented,
    b. the strategy (for e.g. recruitment, motivation) is done with a proven methodology,
    c. there are specific, defined, measurable, acceptable goals for response rate and representativity,
    d. response rates are calculated according to international standards (e.g. AAPOR),
    e. representativity is assessed via R-indicators and/or coefficient of variation,
    f. registration and completion rates for relevant population subgroups are monitored,
    g. proven methodology is used to process collected data to statistical output, and
    h. plausibility checks are specified, and the smart data satisfies these plausibility checks.
3. It is clear what information the used smart solution is providing.
4. Machine learning routines follow literature best practices and result in sufficient quality for the concerning smart solution. The performance of these algorithms satisfies specified thresholds in out-of-sample use (e.g. through accuracy/F1 scores).

5. UI/UX design of the smart solution is usable and works for the specific solution. It is not yet optimized and cannot be shared between solutions. Usability includes:
    a. to attract and retain participants is checked against a) engagement, b) accessibility, c) instructions/information, d) time efficiency, e) feedback and error handling,
    b. willingness to share detailed personal information checked against a) trust and credibility, b) security and privacy, c) transparency, d) data collection efficiency, e) user control, and
    c. ability to complete complex tasks, a) intuitive, b) task flow and guidance, c) error prevention and handling, d) training and onboarding, e) feedback and support.
6. In multi-mode settings:
    a. smart nonresponse and selection differences can be evaluated,
    b. smart measurement differences can be evaluated, and
    c. smart measurement differences can be adjusted for, to warrant comparability in time and between relevant subgroups.

**Business process**

1. The business process was designed and applicable for the production of statistics with the smart solution.
2. Staff is trained and assigned to new roles (roles specific for smart solutions).
3. The process is not standardized. The process is rather ad hoc. The process is probably shaped as a stovepipe for a given survey.
4. There is a process for the data collection, but also for the processing and analyses sub processes.
5. An acceptance test of the entire chain has been carried out.
6. Fieldwork is monitored via a minimum set of indicators based on the process data. Process data is collected; however not per se in a standardized way.
7. Interviewer feedback, if applicable, is evaluated and summarized.
8. The contact centre / helpdesk has the capability to answer questions of respondents regarding the use of the smart solution.
9. App store analytics are performed (downloads, ratings, etc.).
10. App usage traffic measured by responses (data is received by NSI).
11. When changes are made to the process, an impact analysis is first carried out.

**IT**

1. Smart services have been demonstrated to follow input and output specifications as provided by the methodology level (quality metadata) and legal level (PET).
2. The IT architecture and all levels of the IT solution (smart solution, smart feature/service, machine learning algorithms, backend) are described.
3. Stress tests and technical tests for the smart solution(s) have been performed.
4. Process to incorporate changes (including improvements of user experience and usability) and bug-fixes into the smart solution is ad hoc.
5. Process to evaluate new operating systems and devices is ad hoc.
6. Process to update smart solutions for new version of libraries and operating systems is ad hoc.
7. Process to upload and manage smart solutions (e.g. apps) on app stores (like Google Play and Apple's App Store) using a company account is managed centrally.

8. The mainstream devices are available for testing purposes.
9. A testing infrastructure to test smart solutions is available and supported.
10. Deployment strategy has been described explicitly to host microservices and to connect to platforms
11. Pentest has been carried out and critical levels are addressed
12. Interaction with external databases is agreed on a functional, technical and legal level

**Legal**

1. A DPIA (including risk analysis) is available for the used smart solutions and assessed by a legal officer. The process complies with the applicable rules and legislation. Thereby the privacy-by-design choices are motivated, documented and approved.
2. The privacy-enhancing-techniques (PET) applied (privacy-by-design) are described.
3. In-house monitoring and handling of smart data errors have been motivated, documented and seconded by legal officers
4. Risks (e.g. security) in smart solution(s) have been assessed, evaluated and sufficiently mitigated according to the relevant officers (e.g. chief security officer).
5. Respondent data control has been tested and evaluated.
6. Informed consent procedures (like the use of sensors) are tested, optimized and compliant with legal requirements.
7. Internal monitoring and handling of smart data errors are motivated, documented and approved by legal personnel.
8. Respondent data control requests are evaluated and approved.
9. It has been proved that the smart solution does not lead to potential security breaches. For example, by conducting a penetration test.
10. Data collection reports (respondent requests, communications) are checked against PET and discussed with legal officers, but there is not yet a standard process in place.

### 3.3.5 Managed

**General characterization**
Smart surveys are used for several statistics. In contrast to the Production level, this level has uniform services with uniform processes, procedures and systems. The organization is more in control and the outcome is more predictable. The different aspects (Methodology, Process, IT, legal) are well-defined and managed. The organisation is able to produce achievable results, which are well managed. Roles and responsibilities are well defined. The focus is on the entire chain, across all statistics. The pipeline idea is increasingly disappearing. Roles and responsibilities are well defined. The PDCA cycle plays an important role at this level: it is continuously monitored to see how the process performs. Improvements are focused across studies but typically still per main process stage (data collection, processing, analysis).

**Organization**

1. The production department has the lead. The focus is on efficiency and standardizing.
2. Smart solutions are applied for different surveys.
3. Knowledge of smart surveys / solutions is quite widespread. Mainly in the data collection, methodology and IT departments. More and more personnel are involved.

4. The organization uses a dashboard (acceptance criteria and standards) to determine whether they are in control. The content of the dashboard is regularly evaluated. Acceptance criteria and standards can be adjusted. New acceptance criteria and standards may also be added or even dropped if it is judged that it is no longer necessary to monitor the acceptance criteria.
5. The staff knows where to report incidents. The incidents are recorded, structurally analysed, reported and action taken where necessary.

**Methodology**

1. Methodology is proven and effective.
2. Effective methodology is available regardless of the type of smart solution.
3. Quantitative monitoring and analysis of the methodology is done.
4. Relevant summaries of in-device paradata/audit trails for smart services have been defined.

**Business process**

1. Monitoring of app store analytics (downloads, ratings, etc.) is performed regularly.
2. Monitoring of app usage traffic measured by responses (data is received by NSI) has been setup.
3. There is a process in place where machine learning predictions with low classification probabilities (so, input that cannot be classified automatically), can be recognized and handled manual by personnel in the back-office.
4. There is a standardized process in place to update machine learning routines, to preserve the required performance.
5. Helpdesk and other interactions with respondents are evaluated and summarized.
6. The process of conducting pilots is standardized and is consistent with the production process. The pilots also use the available business services that are available for production.

**IT**

1. Process to incorporate changes (including improvements of user experience and usability) and bug-fixes into smart solutions is well-defined and managed.
2. Process to evaluate new operating systems and devices is well-defined and managed.
3. Process to update apps for new version of libraries and operating systems is well-defined and managed. A backlog is maintained on a continuous basis.
4. Monitoring of new versions of libraries used (libraries update frequently) has been setup
5. Monitoring of new version of operating systems has been setup.

**Legal**

1. Risks (e.g. security) are regularly re-evaluated in terms of prevalence, likelihood and impact and discussed with security officers. Action is taken where necessary.
2. Data collection reports (respondent requests, communications) are checked against PET and discussed with legal officers, through a standardized process.

### 3.3.6 Optimized

**General characterization**

This last maturity level is characterized by being able to respond quickly and proactively to changes/changing needs in order to remain the best. The organisation is a centre of knowledge and excellence. Quality and efficiency are beyond question and are seen as a key differentiator. The organization is increasingly focusing on the external world and on partnerships. Customers play a prominent role in this, just like all other chain partners. The organization is continuously looking for opportunities to position itself even more strongly as a leader in the field of smart solutions for statistics. Internally, the organization is so well established that it can respond quickly to changes. The organization is a learning organization. All the aspects (methodology, process, IT and legal) are continuously improved. Continuous improvement is embedded in all of the processes. It is in the veins of the organisation. Employees are empowered to take action and innovate. New innovations can be applied easily. The PDCA is focussed on still becoming better and more efficient and on the statistical business process as a whole.

**Organization**

1. Several departments play an important role in this level: Research and Development, Account Management and Human Research. The innovation/Research and Development team has a strong position in the organization. The organization continuously looks for new ideas in the outside world.
2. The organization maintains good contacts with stakeholders: customers, but also all other parties in the supply chain. The various parties work together to organize the chain as efficiently as possible.
3. The organization is a learning organization. The organization strives for improvement, innovation and adapts to what is changing around it. There is a culture of knowledge sharing across the organization. All concerning smart solutions.
4. The organization is focused on identifying changing capacity needs (resources, skills, knowledge), focusing on changes in future developments concerning smart solutions.
5. Teams are empowered to propose and implement improvements.
6. The organization enters into partnerships with data source suppliers (e.g., energy companies) from the rest of the chain. The entire chain is optimized and not just your own process.

**Methodology**

1. Smart method effects are evaluated at least every five years.
2. There is a capability to generate and develop original ideas independently.
3. Methodologists have the ability to search for and explore new ideas and trends in the external environment.

**Business process**

1. Processes are not only defined and managed but are continually improved based on data and feedback.
2. Performance of the process is monitored using quantitative metrics and data-driven techniques.
3. Processes are flexible and can be adjusted quickly to meet changing business needs or external conditions.

4. There is a business process in place for collaboration with external parties, with clearly defined responsibilities.
5. Advanced risk management practices are integrated into the process.

**IT**

1. In-app usage monitoring (on approval) for improving user experience and usability.
2. Performance of the IT solutions are monitored using quantitative metrics and data-driven techniques.
3. The organization is continuously searching for and implementing state of the art IT solutions. New ‘smart’ features can be implemented quickly in existing smart solutions or new smart solutions.

**Legal**

1. Demonstrates expertise in establishing comprehensive legal frameworks and strategically identifying and evaluating available options.
2. Skill in performing risk analysis and identifying which significant risks need further examination.

# 4. Business functions and Maturity model

## 4.1 Introduction

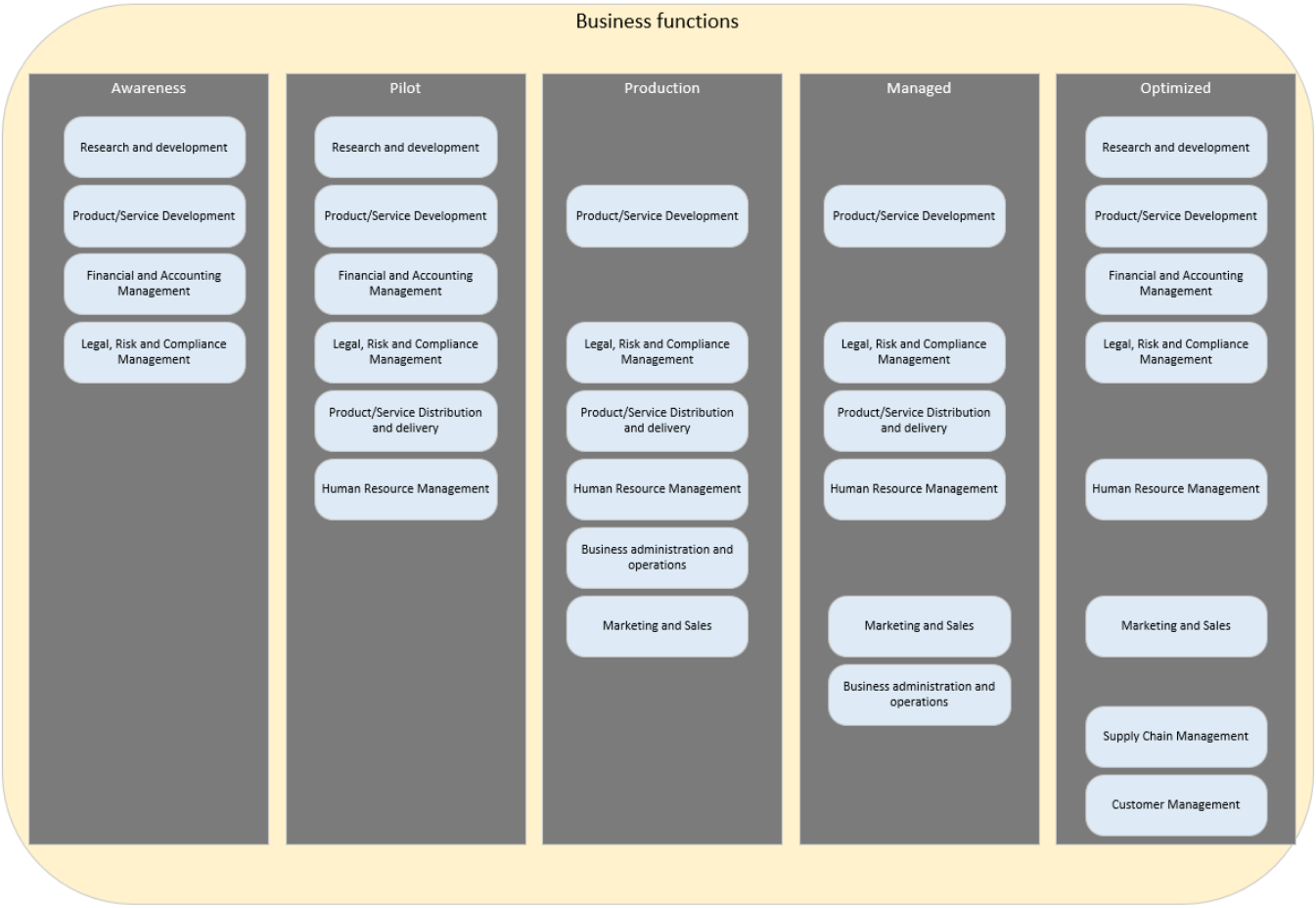
A business function[14] focuses on what the organization must be able to do, to create its products and services. It achieves a specific purpose. It's about a description of the what, not the how and with what. Examples of business functions are:

- Research and Development
- Product/Service Development
- Product/Service Distribution and delivery
- Marketing and Sales
- Customer Management
- Business Administration and Operations
- Human Resource Management
- Financial and Accounting Management
- Supply Chain Management
- Legal, Risk and Compliance Management

Each level of the maturity model has different business functions that are primarily involved. This chapter describes the relationship between business functions and maturity levels. The business functions listed are example functions. Depending on the situation, you may or may not use the function.

[14] Sometimes also called 'capability'

## 4.2 Model




### 4.2.1 Maturity level: Awareness

The business function that plays the most prominent role at this level is Research and Development. The overview below contains a list of functions that you may have to deal with at this level:

- Research and development
  - Consideration to do something with smart surveys, e.g. an app.
  - Determination of what is essential for smart surveys and what is still missing in the NSI (first estimate), such as technology, knowledge, etc.
  - Identification of stakeholders
  - Decision-making (go, no go) to carry out pilots
- Product/Service Development
  - Determination possibilities to use smart surveys for which statistics (first estimate).
- Financial and Accounting Management
  - Elaboration of a Business Case (high level, in any case the benefits), including assessment of whether it meets the objectives of the NSI.
- Legal, Risk and Compliance Management
  - Establishment of legal frameworks and options (high level).

### 4.2.2 Maturity level: Pilot

Business functions that play the most important role in this level are Research and Development and Product/Service Development. A list of relevant functions can be found below:

- Research and development
    - Determination of what is essential for smart surveys and what is still missing in the NSI (in detail), such as technology, knowledge, etc. Arranging what is needed and what is missing.
    - The organization is convinced of the smart solution.
    - Creation and implementation of a communication plan for stakeholders within and outside the organization.
- Product/Service Development
    - Preparation of acceptance criteria (in advance) (preferably with standards determined), so that they can assist in the decision-making process about whether or not to implement the solution in production.
    - Execution of the pilot by the project team. This work is often one-off or temporary in nature.
    - Preparation for respondents to be able to use the pilot app.
    - The decision-making (go, no go) to go into production with smart surveys.
- Product/Service Distribution and delivery
    - Placement of app in app store. Requirements towards the app stores have to be fulfilled. And that is a large step. More and more also the functionalities are reviewed by the app store reviewers.
- Financial and Accounting Management
    - Elaboration of a business case in more detail, to determine whether smart survey yields sufficient compared to the costs.
- Legal, Risk and Compliance Management
    - Elaboration of legal frameworks and options (detail) for the implementation of the pilot.
    - Execution of a risk analysis and determination of the greatest risks. Decision what to pilot.
- Human Resource Management
    - Decision whether to outsource app construction or put together an app team.
    - Putting the right people together with the right knowledge.

### 4.2.3 Maturity level: Production

At this level, the business functions Product/Service Development and Business administration and operations play an important role. A list of relevant functions in this level:

- Product/Service Development
    - Decision what type of app to use.
    - Design of (adjustments in) production app.
    - Building and execution of a test (of the adjustments) of a production app.
    - Ad hoc improvement function.
    - Development of an architecture for production process.
    - Design, realization and execution of a test of the app channel, integrated in the production system.
    - Provision and maintenance of work instructions and procedures (high level) for production.
    - Development of training materials for production.
    - Deployment of microservices for production purposes.

- Business administration and operations
    - Preparation operations.
    - Operations.
    - Preparation for respondents to be able to use the production app.
    - Monitoring and analysing.
- Product/Service Distribution and delivery
    - Placement of production app in app store.
- Marketing and Sales
    - Search for new customers interested in smart surveys.
- Human Resource Management
    - Organization set up for production, often with stovepipes.
    - Coordination function between different parties, as a chain.
- Legal, Risk and Compliance Management
    - Elaboration of legal frameworks (in detail) for the entire chain.
    - Performance of risk management per process.

### 4.2.4 Maturity level: Managed

At this level, the business functions Product/Service Development and Business administration and operations play an important role. A list of relevant functions in this level:

- Product/Service Development
    - Decision about which type of app the entire organization will use.
    - Design of (adjustments) an organization-wide app.
    - Building and execution of a test (of the adjustments) of an organization-wide app.
    - Decision (adjustment) organization-wide collection and communication strategy through app.
    - Design, realization and execution of a test of the organization-wide app channel, integrated in the production system.
    - Deployment of microservices for pilot purposes in a generic way organization wide.
    - Development of training materials for services that can be used organization wide.
    - Provision and maintenance of work instructions of services that can be used organization wide.
    - Standardization of pilot process, connected to the production line/system as much as possible, with decision-making based on predefined acceptance criteria and standards.
    - Working in accordance with architecture, focused on generic services.
    - Establishment of organization-wide policy.
- Business administration and operations
    - Preparation operations.
    - Operations.
    - Preparation for respondents to be able to use the organization-wide app.
    - Monitoring and analysing.
- Product/Service Distribution and delivery
    - Placement of organization-wide app in app store.

- Human Resource Management
    - Standardization of the organization especially for services.

- Legal, Risk and Compliance Management
    - Standardization of legal frameworks (instead of a DPIA per process, a generic DPIA).
    - Performance of risk management organization wide.

### 4.2.5 Maturity level: Optimized

Several business functions play an important role at this level, such as Research and Development, Customer Management, Human Research Management and Supply Chain Management:

- Research and Development
    - Search for new ideas in the outside world.
    - Development of new ideas (own ideas).
- Product/Service Development
    - Application of the new successful ideas to (other) products/services.
- Marketing and Sales
    - Selling the new ideas to (potential) customers.
- Customer Management
    - Proactive approach and influencing customers. The identification of the changing needs of customers.
- Human Resource Management
    - Strong positioning of Innovation team.
    - Design of a learning organization.
    - Identification of capacity needs (looking ahead) aimed at changes in future developments.
- Financial and Accounting Management
    - Elaboration of a Business Case, including assessment of whether it meets the objectives of the NSI.
- Supply Chain Management
    - Proactive approach and influencing stakeholders.
    - Entering into partnerships.
- Legal, Risk and Compliance Management
    - Establishment of legal frameworks and options.
    - Carrying out risk analysis and determination of which major risks require further investigation.

# 5. Process Building Blocks

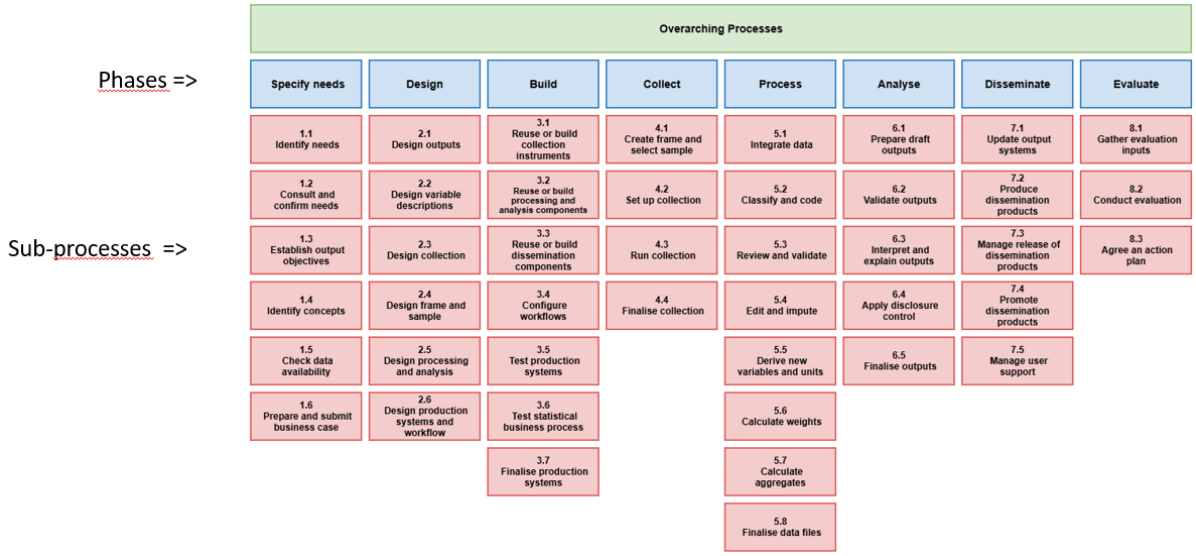
## 5.1 Introduction

The main objective of WP4 is to deliver guidelines that will help NSIs to extend their business process to adopt smart solutions in their surveys.

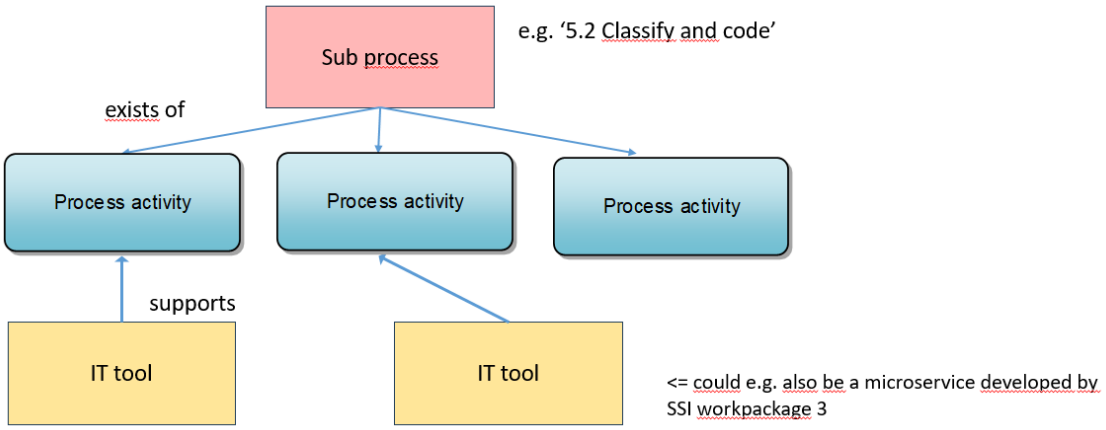
Part of this is to help NSIs to model their business process and to identify the capabilities needed. Each NSI has its own situation, its own context, and its own ideas of applying smart solutions, and thus has its own process requirements. It is not feasible for WP4 to make a process model for each of these situations.

So, what we have developed are **process building blocks**. A building block should be seen as a **business process activity**.

Looking at the GSBPM framework we see phases and sub processes defined.




Each sub-process exists of multiple process activities. And a process activity can be supported by IT-tools/systems.



In our 'library' with building blocks we recognize two types of process activities:

- process activities that are **new**; so **non-existing for non-smart surveys**. E.g. 'Develop app'
- process activities that are **not specific as such for a smart** survey but needs extra attention. E.g. 'Train helpdesk employees'

In GSBPM there is a sub process called "*Design production systems and workflow*". GSBPM states that "*This sub-process determines the workflow from data collection to dissemination, taking an overview of all the processes required within the whole production process and ensuring that they fit together efficiently with no gaps or redundancies […]*" (see reference [6]).

The building blocks will support this activity. The idea is that an NSI can use these building blocks to model their production process. Each NSI can vary in the order of the process activities and determine which blocks appear or do not appear in the process. It can be used to model a generic process or a survey-specific process. That is up to each NSI.

To demonstrate how this works, we have provided a business process model as an **example**. This shows how building blocks can be used to compose a process model.

The building blocks can also just be used as a checklist, which provides an NSI information of topics to address concerning the implementation of smart solutions.

Building blocks are non-NSI specific and have enough detail to show the 'smart' aspects. There are of course many process activities that are needed in a statistical business process, but it is too much and also not of added value to mention them as far as they are not affected with respect to smart solutions. We only looked at building blocks that are relevant for using smart solutions.

The scope is all the GSBPM process phases up to and including the 'Process' phase. For the 'Analyse', 'Disseminate', and 'Evaluate' phases, we do not see extra process activities (in the context of smart solutions). So, no building blocks are provided.

- Some process activities are general usable/applicable, thus **not specific to the type of smart solution**
- Other activities are **specific to the type of solution**. In this respect we restricted ourselves to the three SSI use cases HBS, TUS and Energy Usage use case. E.g. "OCR of the receipt" [HBS]

We related **microservices**[6] – as far as in the scope of SSI – to the building blocks. In this way, an NSI gets a clear picture of where microservices can be used in the business process.

*In deliverable 3.3 from WP3 (see reference [4]) is described that a microservice can use a multiple set of ML-models. E.g., the OCR microservice uses three ML-models*[15]*.*

In addition, for each building block, related **actors** are identified, i.e., for each process activity, it needs to be clear what business actor is involved.

Process activities – in general – tell you *what* to do, they do not say anything about the *how*. So, the building blocks will not address the methodology aspects and will also say nothing about IT solutions (other than the microservices and the ML-models that could be used). Methodology, legal but also IT will provide requirements for your business process. Methodology will also give advice about which

[15] The first model focuses on the 'preparation' of a ticket (contour, orientation), the second model deals with how text from a ticket is extracted and placed into position boxes, and the last model connects the boxed text to standardised labels holding relevant elements that (can) appear on a ticket.

methods work best in certain situations. E.g., there will be a building block that says that you need to ask/inform the respondent about consent. But the building block will not say how to do that.

However, we will – when applicable – mention alternative options[7]. To this end, we also refer to the results of the other SSI work packages.

Based on all this, we then also look at the GSBPM and BREAL framework. We have positioned building blocks within these frameworks. See appendix B. In this way, we elaborated the GSBPM with the aspects that are specific to using smart solutions.

One more remark on the building blocks: Parts of a process are often iterative. This is also what GSBPM says. The building blocks itself have no iterative aspect; the iterative only comes out when you start using the building blocks to create the process.

## 5.2 Building blocks

In this chapter the complete set ('library') of building blocks is described.

Building blocks are combined into groups. E.g., the group 'App'. The idea is that when you, as an NSI, decide to use an app in a survey, you can look at the building blocks in this group. Then these are all candidate activities to include in your process.

*No strict rules were used to create a group. It is just meant to give you some structure. They are not grouped by process phase (like 'design', 'build' or 'collect') but by subject (like 'app' or 'Helpdesk'). Please feel free to group them otherwise to fit your needs.*

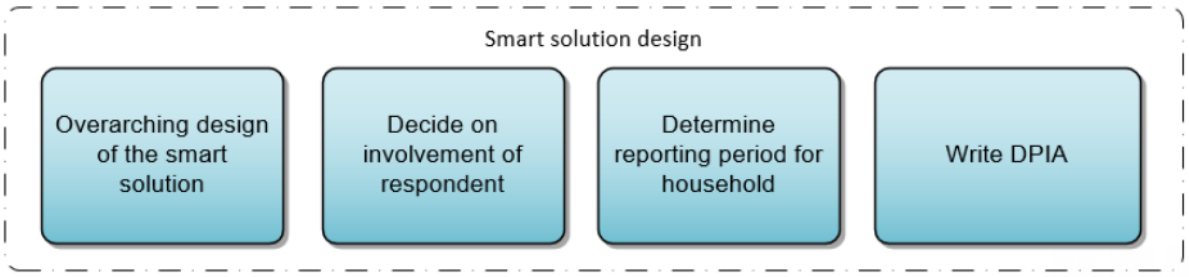
The groups we use are:

| See chapter | Group | Comprises process activities concerning… |
|---|---|---|
| 5.2.1 | Smart solution design | ...designing the overarching design of the smart solution |
| 5.2.2 | App | …designing, building, downloading and installing an app |
| 5.2.3 | Machine-learning model – Training & Maintenance | ...developing and maintaining a machine-learning model |
| 5.2.4 | Collection and communication strategy | …designing strategy and inform respondent about consent and privacy |
| 5.2.5 | Integrate into production system | …the integration of the 'app channel' into the rest of the data collection landscape |
| 5.2.6 | Providing receipt | …scanning / uploading a receipt and entering additional information |
| 5.2.7 | OCR/NLP | …deploying the microservice, OCR/NLP the receipt, involvement of the respondent in this |
| 5.2.8 | Geolocation | …deploying the microservice, collecting geotracking points, derive motion and stops |
| 5.2.9 | Suggester lists | …maintaining shop, product, activity lists |

| | | |
|---|---|---|
| 5.2.10 | Diary | …filling the diary with the individual receipt information, determining the moment of response |
| 5.2.11 | Interviewer | …the role of the interviewer |
| 5.2.12 | Data processing | …processing activities |
| 5.2.13 | Classifying | …classifying to COICOP, HETUS Activity coding list (ACL) |
| 5.2.14 | Helpdesk | …the role of the helpdesk |
| 5.2.15 | Monitoring and analysing | …the monitoring and analysing of the process and the app usage |
| 5.2.16 | Provide respondent with smart device | … the use cases where you want the respondent to use a smart device that is not already in possession of the respondent. E.g. an energy dongle or an activity tracker. |

Everything starts with the question of what smart solution you would like to use in a certain survey. This depends on the needs (see GSBPM 'Identify needs') and on the possibilities (financial, IT, etc.) you have as an organisation. Also, the level of maturity of your organisation regarding applying smart solutions plays a role: to use a smart solution for the production of statistics you need to be at a certain maturity level (see chapter 'Maturity model').

## 5.2 1 Smart solution design




**5.2.1.1 Overarching design of the smart solution**

This is about deciding which smart solution to use and about designing the overall workflow of how the smart solution will be used. It is about designing the respondent journey how the respondent will interact with e.g. the app.

There is an active-passive trade-off decision, i.e. processing in-survey respondent involvement versus processing post-survey. There must be a moment in which decisions are made about what processing to do (near) real-time, what with the help of respondents and what post-survey. It is more than 'just' the app design. It is a higher-level, overarching, design stage where the overall workflow is constructed.

Actor: Survey designer, methodologist, IT architect, privacy officer

#### 5.2.1.2 Decide on involvement of respondent

What type of interaction will the respondent have, concerning the smart solution? What type of information will be fed back to the respondent? Will the respondent have an active or passive role? What processing will be done (near) real-time, what with the help of the respondent and what post-survey?

When using an app, this is an important decision to make. E.g., in HBS there are a few feedback options:

- Feedback on the quality of the photo (when quality is too low: take new photo)
- Results of OCR
- Results of COICOP classification
- Return nothing

And subsequently the question is whether or not and in what way the respondent can make adjustments, additions or can delete items.

If the OCR model doesn't produce near-complete and correct classifications, the already quite heavy response-burden grows even heavier if the respondents have to fill in or correct classifications. This might lead to increased levels of dropping out.

On the other hand, if the respondents are not asked to correct or fill in missing information, they are left with the impression that incorrect and/or inadequate data are going to be used, which will lead to false results being published. This belief might lead to a PR[8] problem for the NSI.

These decisions have a considerable impact on the business process, the app design and the IT architecture.

The classification of a receipt text into a COICOP category is normally an activity in the 'Process' phase (see GSBPM, subprocess 5.2 'Classify and code'). However, returning COICOP results to the respondents requires a more or less real-time feedback. This requires that the COICOP model is available for the app and maybe even *on* the device (see WP3; this is also a microservice).

Maybe you do not want to feed back on the 4$^{th}$ or 5$^{th}$ level of COICOP (because maybe it is too fine grained for the respondent). So, that supposes two 'versions' (our outcomes) of the COICOP algorithm: on the most detailed level (as output for the statistics) and on a higher level (to return to the respondent).

Another example on is the use of a dongle within the Energy Usage use case. During various phases of the survey, different levels of feedback can be introduced:

- In the first phase, use of an appliance is noted at high detail in a diary and energy data is collected from the smart energy meter via a dongle. This enables the NSI to attribute found energy use patterns to appliances.
- In the later phases, the level of detail of the diary can be much less, which helps reduce the burden put on the respondent. The specific model for the household that is surveyed, will attribute the energy use automatically.

Actor: Survey designer, methodologist, IT architect, privacy officer

#### 5.2.1.3 Determine reporting period for household

The requirement in HBS is that the diary should not include expenditures from a period when the household is on holiday. Therefore, before the respondent starts filling in the diary, the holiday period is asked. If the holiday overlaps with the initial reporting period, the reporting period needs to be changed.

The activity is mentioned here because a choice should be made whether the holiday question is within the app or not. This has an impact on the design of the app but also on the process flow.

Above example also applies to the Energy Usage use case.

Actor: Survey designer

#### 5.2.1.4 Write a DPIA

The carrying out of a DPIA for smart surveys should be considered mandatory. The General Data Protection Regulation (GDPR) requires a data protection impact assessment to be carried out when the processing may present a "high risk to the rights and freedoms of natural persons". DPIA is particularly important when new data processing technology is introduced. See reference [7].

The DPIA should assess the privacy impact of smart solution not "in the abstract", but in the specific context of the statistical survey in which they apply.

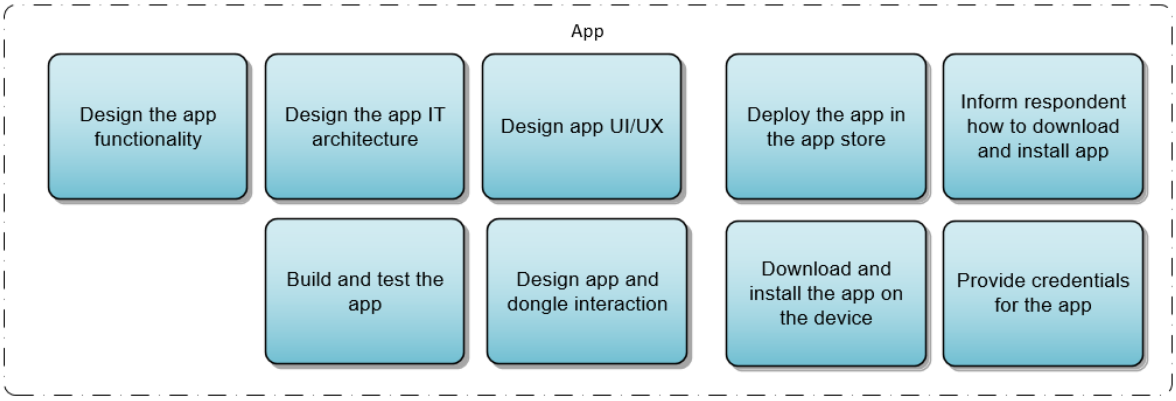
Examination of the experiences of the NSIs in the various countries (see deliverable 5.2 of SSI WP5) shows that separate DPIAs are usually carried out for *standard* processing operations with common characteristics (e.g. processing operations concerning current surveys using the same techniques and methods of data collection and processing) and for *non-standard* processes e.g. when new techniques and methods are introduced into a standard processing operation. The implementation of smart solutions should be considered as non-standard processing and thus needs a separate DPIA.

SSI WP5 provides a DPIA template for this (see reference [7]). The template applies to the case where new techniques and methods (the "smart features") are introduced within a statistical survey already planned and currently carried out by the NSIs of the ESS (e.g. HBS or TUS), which therefore has a predefined statistical purpose and legal basis. The two-step approach described and proposed in the SSI (see WP5) leads to a motivated and clearly defined set of privacy-by-design and data-minimisation choices for each type of smart survey. These choices allow for balancing on the basis of real and empirical data provided by the respondents.

Using an external third party in your business process (e.g. for developing an app) is a complicating factor. This leads to additional requirements, contracts, data flows, etc. and will impact privacy and security aspects. This will require extra attention in writing your DPIA.

Actor: legal officer, project manager, chief security officer

### 5.2.2 App




When you decide to use an app for your smart solution (e.g. in HBS or TUS), the building blocks in this paragraph are relevant.

GSBPM has a 'Design' and a 'Build' phase. The same two phases are relevant for developing an app. Within 'Design' we make a distinction between a) the functionality b) the IT architecture and c) the UI/UX.

#### 5.2.2.1 Design the app functionality

Resulting from the overall survey design there are needs (requirements) for the app. So, the main purpose of this process activity is to 'translate' these requirements into a design of the app. Which functionality should the app provide and what tasks should the respondent be able to do within the app? In case of HBS: are you e.g. also going to provide a function to upload an e-receipt?

This should include sufficient attention for legal and security risks. 'Privacy-by-design' is a good principal to follow. The question is e.g. what personal and identifying data will be available through the app.

> Privacy by design is a data privacy concept that calls for the incorporation of data privacy protections into the design of information systems, products, and services. The goal is to prevent data privacy breaches and protect the privacy of individuals by proactively incorporating data privacy safeguards into systems and processes.

In the WP5 deliverable (see reference [7]) there is a taxonomy of types of apps, with some characteristics mentioned (see reference [7] for more information):

| App Type | Platform Compatibility | Development Cost | Performance | User Experience | Access to Device Features |
|---|---|---|---|---|---|
| **Native App** | iOS, Android | High | Very High | Excellent | Full |
| **Web App** | Any (via web browser) | Low | Low | Fair | Limited |
| **Hybrid App** | iOS, Android | Medium | Medium | Good | Moderate |
| **Progressive Web App** | Any (via web browser) | Low | Medium | Good | Moderate |

You should consider what type(s) of app you would like to use and what would be viable in your survey environment.

Having an external audit on the app before publishing it, is a must. It should also be noted in planning the app and the development timeline that the NSI's policy might require a security audit on the app and the environment it connects to, for which time and resources should be allocated, including implementing possible changes.

Machine-learning models may be part of the app. E.g. in HBS for the OCR and for the COICOP classification, and in TUS for labelling activities and in the Energy Usage use case for usage patterns. This should be incorporated in the functionality of the app.

Actor: App designer, Methodologist, Process architect, IT architect, Privacy officer

#### 5.2.2.2 Design the app IT architecture

You need to decide whether you are going to use a designated app or a web-based app, or both. This choice is often determined by the need to access specific device-specific features.

In addition, several policy choices will also have to be made, such as which types of smartphones need to be supported, operating systems, web browsers, etc. Also, a separate version for Android and iOS is needed. The two app stores have different requirements for the app, and there are some differences in functionality as well.

Part of the IT design of the app is how the app will communicate with the back end and the data collection platform. Also, performance ('speed' of the app) and other non-functional requirements need to be addressed. The app e.g. also needs to follow a long list of requirements (stated by Google/Apple) in order to be published.

Actor: IT architect, app designer

In reference [3], we read: *"When administering smart surveys, researchers have the choice to do so via a designated survey app that participants need to download to their smartphone or let participants complete the survey via a mobile web browser. A variant of the web-based approach is so called (progressive) web apps, which basically mimic the look and feel of a full-blown app but still open in the (mobile) browser of the participants device. [...]*

*A designated app allows a better integration of smart elements into the survey, taking advantage of the full range of smart features of a smartphone (e.g., sensors, camera). Apps also can operate for longer time periods without an active internet connection. Finally, they do allow to actively contact the participants via messages directly from within the application, for example, by sending push notifications to remind participants to complete the diary. This feature makes apps very attractive for intensive longitudinal data collection, such as TUS and HBS.*

*Based on the few findings in the literature, we [SSI] suggest using an app for data collection in TUS and HBS, which allows a seamless integration of smart survey features into the survey and more direct interaction directly through the app with the participant throughout the data collection period (assuming days of data collection >1). In addition, NSI will need to decide whether they offer participants a browser-based alternative to be more inclusive and allow people who do not want to or cannot download an app to participate. Ideally, this would be done with a progressive web app, to reduce potential measurement error due to different instruments being used."*

#### 5.2.2.3 Design app UI/UX

Based on the business and IT requirements, you need to design how the UI/UX should look like. The design is influenced by the app platform; iOS has e.g. different design guidelines and standards than Android.

Actor: UI/UX-designer; app designer

#### 5.2.2.4 Design app and dongle interaction

This kind of smart solution requires a next level knowledge on the technical protocols to be used to interact between dongle, app and platform.

Actor: App designer, IT-architect

#### 5.2.2.5 Build and test the app

Based on the design the app should be built. Also, functionality and usability testing are part of this activity.

In the SSI-cases concerning HBS and TUS, the microservices and machine-learning models are part of the app. They have to be integrated into the app.

Some NSIs will build the app themselves (in-house), some will use an external supplier, and others will (re)use an existing app. Another option is for NSIs (and/or other parties) to jointly build and maintain the app.

In the first case (build in-house), an NSI should have app developers and app testers available. Also, you need specific app development tools and different sorts of test devices. This can have quite an impact on the organisation. What your NSI needs and how you organize this, depends on the level of maturity (see chapter 'Maturity model').

Regardless of the choice your NSI makes, you need to think about the governance of your app management. The type of governance required depends on the relevant option as mentioned above, and on the other hand also depends on the maturity level of the NSI (see SSI Maturity model).

Actors: App developer, app tester, application manager, security officer, group of 'test'-respondents (people who will actually use the app for testing purpose).

#### 5.2.2.6 Deploy the app in the app store

The developed app should be made available at a location where the app can be downloaded by the user. For Android and iOS, there is a different app store.

There is an important legal/privacy aspect, as you have to fill in the Google and Apple forms.

Requirements towards the app stores have to be fulfilled. And that is a large step. More and more also the functionalities are reviewed by the app store reviewers.

You should also think of managing the versions of the app in the app store and the process for updating the apps already installed on the devices of the users. E.g., when there is a new version, will you also keep the old version (or are apps on the devices updated mandatory)? Should a new version only be put in the store between two collection cycles? Keep in mind that you are dependent on the app store review process. That process can take a considerable amount of time. And how do you deal with bugs? Do you address critical bugs only after the end of a collection cycle or can you release new versions in between? And how does that affect your final result? These questions do not have an obvious answer and should be thoroughly thought through.

A note: the name of the app should be thought out carefully, especially in countries where more than one language is used in the survey.

The app can be published only regionally, but as the number of countries providing an app increases, it should be noted that every HBS app can't be named "HBS App", since it might lead to respondents accidentally downloading the app of another country's NSI. Because of this, the name of the correct publisher (NSI) should be included in the respondent instructions.

Actor: App developer

#### 5.2.2.7 Inform respondents how to download and install the app

The aim is to inform the respondent where/how to download and install the app. This can be done, for example, in the letter of invitation. Also, an interviewer could play a role in this (see chapter 'Interviewer'). And what information do you provide in the app store?

Actor: Communication officer

#### 5.2.2.8 Download and install the app on the device

The aim of this activity is for the respondent to download the app and install it on their device. The trigger for the respondent will be the communication from the activity 'Inform respondent how to download and install the app'.

Actor: Respondent, possibly with the help of the helpdesk or interviewer

#### 5.2.2.9 Provide credentials for the app

The premise is that if the respondent is going to use the app, he has to authenticate (and e.g. activate his account). To do so, he will need to have credentials.

If there is a web questionnaire in addition to the app, the NSI will have to determine whether the credentials are the same or specific for the app. When using an existing integrated platform, you are bound to the functionality of the platform; and then it is probably no issue.

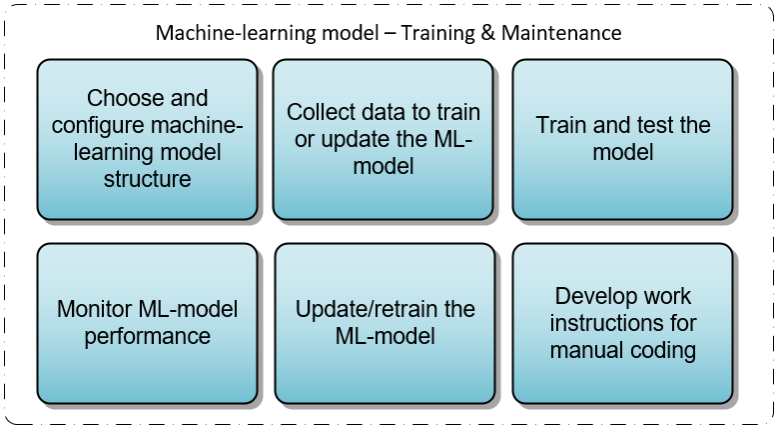
An important point is to check that the link between the credentials and different modes of collection, are well defined, so that the data can be integrated.

In the case of household surveys, the NSI will also need to determine whether the credentials are for the entire household, or whether there are separate credentials for each person in the household. In the latter case, though, the process will be more complex.

There is a privacy issue, and the choice also depends on your DPIA.

Actor: Communication officer, privacy officer, IT architect

### 5.2.3 Machine-learning model – Training & Maintenance




This group of building blocks is relevant when you are going to develop and maintain a machine-learning model (e.g. for OCR/NLP, Geolocation or Energy Usage patterns) in your process. These activities are part of the PDCA-cycle for training and maintaining the model (see chapter 'PDCA-cycle for machine-learning models in smart surveys').

*For developments and considerations on methods for processing sensor data in the contexts of receipt scanning and geolocation, please see deliverable 2.2 of WP2.*

*NLP = Natural language processing.*

*There is a distinction between OCR and NLP. While **OCR** focuses on extracting text from images and converting it into characters, it does not inherently understand the meaning or context of the text. **NLP**, on the other hand, is concerned with the comprehension and analysis of language semantics and can process and interpret the extracted text from OCR or other sources. In some applications, OCR and NLP can be used together. For example, after OCR extracts text from an image, NLP can be applied to understand the meaning of the text, extract relevant information, or perform additional language-related tasks. Also, the term "**Document Understanding**" is sometimes used. This is a specialized application of NLP focused on extracting and interpreting information from documents.*

In case of TUS, the geolocation model is used to derive motions and stops from geolocation points.

In case of HBS, the OCR/NLP model is used to extract text from receipts and to derive the correct metadata (E.g., is it a shop name? Is it a product name? Or product price?). In HBS also a ML-model can be used to classify to COICOP. (See also reference [4] for the models used in HBS and TUS.)

In case of Energy Usage use case, the attribution model of the specific household itself, is only used once. However, the deduction model that enables the NSI to determine what appliances are used when some of them are used in parallel, needs training.

#### 5.2.3.1 Choose and configurate machine-learning model structure

In machine learning, you have to distinguish between the model structure and the rules that a model learns. The model structure is often generic across different tasks/problems and depends on the type of model. For example, you have models based on a decision tree (or multiple decision trees) (such as decision trees and random forest models) and models that work with weights (linear regression, logistic regression and neural networks).

So, first you need to choose a particular flavour of model with a particular model structure. Within that structure you can make (limited) variations; for example, with random forest you can choose the number of decision trees, or the depth of those decision trees. Within the established structure of the model type, you have thus, some freedom to vary the model. This is what we call here configuring a model.

Actor: ML-specialist, Data scientist

#### 5.2.3.2 Collect data to train or update the ML-model

For training the ML-model you need to have data as is input. When first time training the model, you don't have data available from the actual survey. But, when you already have the model used in production, you can use actual data collected from the survey (e.g. receipts from respondents) (see building block 'Update/retrain the ML-model').

The type of input will determine the number of tasks to be carried out for this activity. Getting datasets from e.g. shops, needs other tasks than using OCR data. For some types of information this activity may not be feasible, because the process to make the data available, is too complex and/or expensive.

Candidate input in case of the OCR/NLP-model:

- Receipts gathered during the survey
- CPI scanner transaction data
- Datasets/receipts obtained from shops.

Also, insight in the things that the respondent edited, could be very helpful. E.g. - in case of TUS – when the respondent changed the activity presented by the model. In this case your application should be able to store and give insight in those edits.

In case of HBS, if your NSI decides[9] to use photos of receipts and/or e-receipts to update the OCR/NLP model, an activity is needed to provide them. This means that photos/e-receipts should be stored somewhere so that they are accessible for the 'update the ML-model' activity (see below).

Actor: Data scientist

#### 5.2.3.3 Train and test the model

The goal of this activity is to first-time train the model. For example, via a supervised or unsupervised method. See PDCA cycle for ML-models for more information. The model must be trained so that it provides sufficient quality. Determining how good the quality should be is part of defining the objectives of the model and should be determined up front, in the Plan-stage of the PDCA-cycle (see chapter "PDCA-cycle for machine-learning models in smart surveys").

Getting to the final model - that will be used in production - requires an iterative process of training and testing.

In case of HBS: There are a lot of types of stores, and thus receipts. Is your model only trained on e.g. supermarkets? You need to have a clear vision on that. Are e.g. 80% of all expenses related to supermarkets? Then perhaps you could be satisfied with not using receipts from other shop types. However, you need to have insight in the consequences for the quality and for the respondent interaction.

Actor: Data scientist

#### 5.2.3.4 Monitor ML-model performance

*This activity is part of the "Check"-stage of the PDCA-cycle (see chapter "PDCA-cycle for machine-learning models in smart surveys").*

A machine learning model used in production should be monitored on performance.

The term 'performance' has two sides.

- **Quality**: Is the quality (of the output of the model) still in accordance with the requirements/objectives? Over time, models will encounter input that they do not see during model training. The quality of the results of the model can therefore degrade over time and influence the results of the survey. The number of unlabelled or wrongly labelled items increases. For products (HBS) this is probably more relevant than for activities (TUS).
- **Efficiency**: Is the processing speed of the model still sufficient? The efficiency can degrade e.g. if the number of different products grows. If you want to feed the model results real-time back to the respondent, there are higher demands on the efficiency of the model.

Monitoring should detect model degradation on time. When the quality and/or the efficiency is below a certain threshold the model needs to be updated/retrained (see below).

You need to decide where to give this activity a place in the business process:

- Perform the activity as part of your production process. Then, the monitoring is actually continuous.
- Perform the activity as part of a separate PDCA-process. In this case the monitoring is at regular intervals, e.g. once every 6 months[10]. However, you need the input from your production process.

Actor: Data scientist

#### 5.2.3.5 Update/retrain the ML-model

Training a model is not a onetime effort. It should be maintained/retrained. So, if the result from the monitoring is that the performance of the ML-model does not conform to the objectives/requirements, the model needs to be updated/retrained.

This may mean that you need to collect additional data, for instance, for a store that is missing or for a COICOP category that is underperforming. See the building block 'Collect data to train or update the ML-model'.

'Retraining' or 'Active learning' are two methods that could be used.

> ***Retraining*** *refers to the process of updating a machine learning model using new data, typically to adapt to changes in the underlying patterns or to improve overall performance.*
>
> ***Active learning*** *is a process where the model actively selects specific instances from an unlabelled dataset for annotation (labelling) by an oracle (usually a human). The goal is to improve the model's performance with minimal labelling effort.*

> From reference [3]: *"… to improve the accuracy of the ML models, human interventions (respondent, coder) must be envisaged to assign correct labels. The new labelled item is used to retrain the model to make it more up to date. Over time, the machine learns from humans and becomes more and more accurate.*

*Furthermore, ML methods require continuous updating. Updating can be done fully automated through online learning or semi-automated through active learning. Retraining is ideally done based on incoming datasets while preserving the privacy of the respondents. In practice, when respondents provide data for which processing performance falls below specified thresholds, then this data should be used for retraining ML model.*

*Active learning is the subset of ML in which a learning algorithm can query a user interactively to label data to obtain the desired outputs. In active learning, the algorithm selects the subset of examples to be labelled from a set of unlabelled data. This algorithm represents a key component in Human-in-the-Loop where human and machine intelligence combine to create more accurate models. The problem in the use of ML in a survey then becomes: 1) Build the automation part, 2) Design a mechanism whereby machine alerts human when it needs input and 3) Design an efficient UI to facilitate human machine interaction (Benedikt et al., 2020)."*

Actor: Data scientist

#### 5.2.3.6 Develop work instructions for manual coding

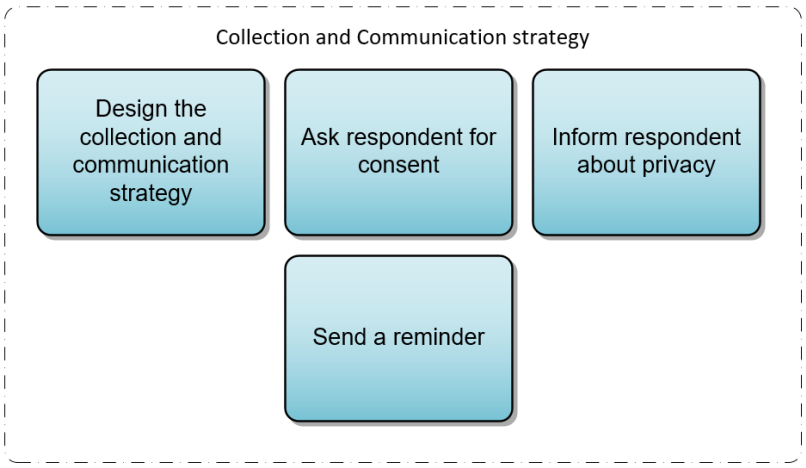
If in your process you will have the activity 'Classify manually' (see chapter 'Classifying (to COICOP / Activity coding list)') this activity is relevant. The classification model should go hand in hand with instructions for the coders. Coders should know how to classify. They need rules and guidelines to reach the required quality.

These working instructions need also to be updated when the model is updated.

*Note: in principle work instructions are necessary for each (manual) process activity. It is mentioned here explicitly because it should not be forgotten in case of updating the ML model.*

Actor: Coding expert, Data scientist

### 5.2.4 Collection and communication strategy




These are all activities concerning the collecting strategy and also the communication strategy, mainly the activities for reminding respondents. Using an app, or applying a smart solution, offers some new opportunities but gives also the need for extra requirements for the communication.

#### 5.2.4.1 Design the collection and communication strategy

Using an app for data collection will have a profound impact on your collection and communication strategy. Questions that need to be addressed are:

- How does the app relate to other modes and questionnaires in the survey?
- How to guide/steer the respondent between the app and questionnaires?
- What will your communication strategy look like? How to inform the respondent about the app and consent? What about the role of an interviewer? How to use reminders? Etc.

Actor: Survey designer, Communication specialist

#### 5.2.4.2 Ask the respondent for consent

Using an app will give rise to paying extra attention to privacy but also asking for consent from the respondent. How to ask the respondent for consent also depends on the legal framework in your country. Some options (so not exhaustively) for asking for consent are.

- Ask for consent through a written consent
- Ask for consent when the respondent first time uses the app or smart feature
- Ask for consent when the respondent sends the data (as soon as done e.g. with the diary)

As an NSI, you need to decide which way to go.

#### 5.2.4.3 Inform respondents about privacy

Privacy concerns may prevent individuals from participating in data collection. So, it is good to give respondents clear and transparent but at the same time concise information regarding privacy. This is of course also the case when not using smart solutions, but in the case of using an app, this needs extra attention. E.g. respondents might get the idea that the app collects all kinds of information, may be even a ‘big brother is watching you’ idea.

Some options are:

- Put the information in the invitation letter
- Put the information in a separate letter
- Show the information in the app or app store
- As a task of an interviewer
- Show the information on the NSIs website

'Privacy by design' is a good principle to use in designing your business process. In fact, it is a GDPR requirement.

**5.2.4.4 Send a reminder**

When using an app, your reminder strategy may look different. In other modes usually the trigger to send a reminder is when there is no response. When using an app, a respondent is often spending extended periods of time collecting data. E.g. 14 days collecting activities or expenditures.
In this case you might consider sending a reminder during this reporting period if you see that the respondent is no longer 'active' (so, stopped recording data).

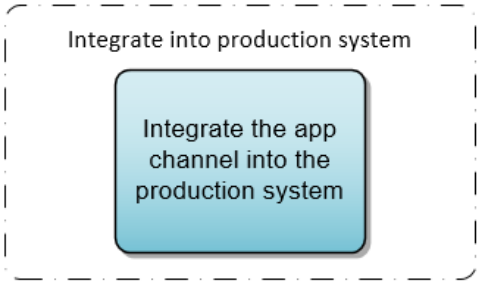
It not feasible to see whether or not the respondent has installed the app, and to use that as a trigger to send a reminder.
Considerations for sending a reminder may include:

- whether the respondent has logged in to the app,
- whether the respondent is still 'active' using the app,
- whether the respondent has recorded data for all days in the diary.

An app provides you also with an additional channel to remind the respondent. You could choose to show a reminder to the respondent in the app / on the device, triggered by the app. E.g. in case the diary is just partially filled. Of course, this only works when the respondent launches the app at all.

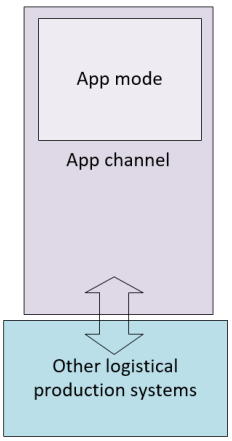
This presumes that the app knows when it's necessary to show the reminder. This also gives requirements for the design of the app.

### 5.2.5 Integrate into a production system




#### 5.2.5.1 Integrate the app channel into the production system

Data collection through an app is a separate (new) mode. The goal of this activity is to integrate the systems that are needed for supporting the app mode, with other (probably already existing) systems supporting the logistics (e.g., your case management platform). The app in itself is just one part of the applications needed to handle the app mode. There will e.g. be backend applications to handle and store the app data and make the data available for the processing phase. To indicate that it is more than just the app, we use the term 'app channel'.



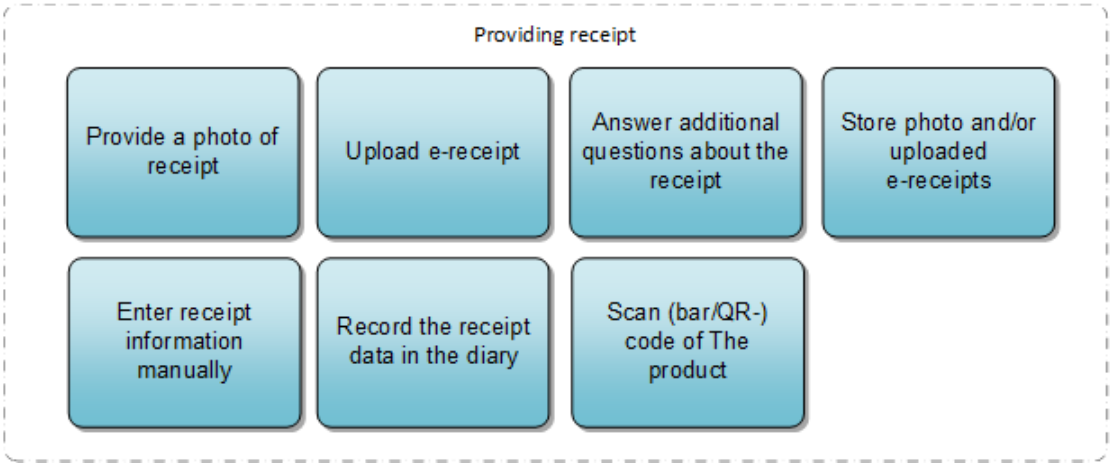

The level of integration depends on the maturity level (see SSI maturity model). When you are e.g. at the third level ('production') you are likely to spend more effort to integrate than in the first and second maturity levels. Actually, at the first level, there is probably no desire/need at all to integrate.

If you have 'stovepipe' processes and systems (process and systems per study/ group of studies), you may need to do the 'integrate' activity more than once. If you have a generic process, you probably also want to have a generic 'app channel'; and then it is probably (hopefully) a one-time effort to integrate.

On the other hand, when you use an 'out of the box' tool/platform that supports (larger parts of) the production system (like MOTUS) and that already has an integrated 'app channel', you do not need to integrate it yourself. This activity has then been done by the supplier. Next, though, the application must be set up and parameterized according to the wishes of the NSI. With MOTUS, the NSI will have its own back-office version to configure this.

Actors: Business and IT architect, process engineer, system engineer

### 5.2.6 Providing receipt (in case of HBS)




The building blocks in this paragraph are all relevant for scanning/uploading a receipt and entering additional information.

#### 5.2.6.1 Provide a photo of receipt

In this activity, the respondent uses the photo camera on his device. The goal is to obtain a photo of a receipt good enough to OCR. That includes e.g. taking the photo, detecting contrast levels, determining the contour of the receipt, cropping the receipt, removing the background, checking the quality of the photo, etc.

Having the complete receipt is important because it contains more info than only the product/service rows. So, no cropping to e.g. only the product rows.

Actor: Respondent

#### 5.2.6.2 Upload e-receipt

This activity allows the respondent to upload an e-receipt (instead of providing a photo). The prerequisite of course is that the respondent must have access to the file.

The file should be in a format suitable for OCR.

Actor: Respondent

#### 5.2.6.3 Answer additional questions about the receipt

The goal of this activity is to ask the respondent to answer some questions about the photographed or uploaded receipt (e.g., the total price). A reason could be that – when known upfront – the OCR model lacks some quality and is not able to interpret some parts of the information. Another reason could be that the output of this activity is very interesting for internal quality checks in the OCR process. Information retrieved from the user could be used as a verification step. E.g., the total price as answered by the respondent should match the total price as derived from the image.

Depending on the performance and quality of the OCR model, the number of respondent actions needed, might change in time.

Actor: Respondent

#### 5.2.6.4 Enter receipt information manually

Instead of taking a photo or uploading an e-receipt, you could also provide the feature that a respondent can enter all receipt information manually in the diary. Providing the respondent with

good quality lists with shops and products that can easily be searched, will have much added value. Instead of freely writing down their purchases, the respondents can be guided via auto-completion lists, which correspond to an implicit coding of the products purchased and the stores.

This activity is not really part of the 'smart solution', but we mention it here because there is a clear relationship with the receipts obtained through the photo or upload.

Actor: Respondent

#### 5.2.6.5 Store photos and/or uploaded e-receipts

The photos and uploaded e-receipts should be stored so that they can be made available for manually coding and/or the analysing process (see chapter 'Monitoring & Analysing') and/or training the ML-model.

There is also a privacy/security issue here to address. Because how long should and may the photos be kept? You should also think about a 'data clean-up / destroy' activity.

Actor: system

#### 5.2.6.6 Scan (bar/QR-) code of the product

The goal of this activity is that the respondent can scan the product code on the product itself[11]. Based on the code the product will be added to the receipt/diary. This presumes that there is a database with products. If in this database these products are also linked to COICOP, that would be an advantage because then the COICOP-code will automatically be available for the scanned product. (this database could e.g. be available based on CPI data)

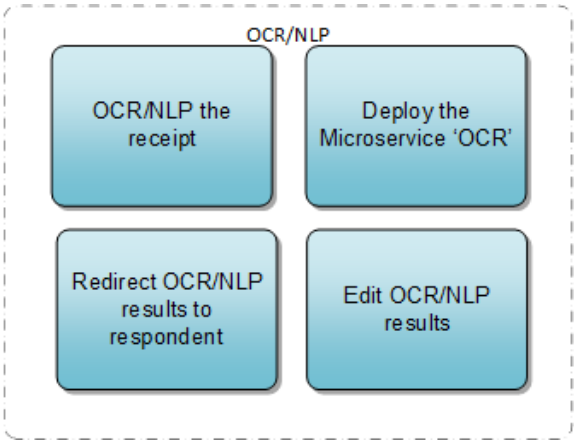
Actor: Respondent

#### 5.2.6.7 Record the receipt data in the diary

The data of each receipt should be part of the diary. This activity adds the receipt data (obtained from the photo, the upload or the manual entry) to the diary. The trigger would be that the respondent indicates that he is finished with the receipt.

Actor: Respondent and system

### 5.2.7 OCR/NLP (in case of HBS)




This group of building blocks is relevant when you are going to use OCR/NLP in your process.

#### 5.2.7.1 Deploy the microservice 'OCR'

The **microservices 'OCR'** is provided by the SSI project. As an NSI you are encouraged to use this microservice.

If you are going to use a microservice it needs to be deployed to the platform to be used. You could also choose to make the microservice part of the app (this is choice made in SSI); in that case the microservice is automatically deployed together with the app.

Actor: IT architect, IT developer

#### 5.2.7.2 OCR/NLP the receipt

The goal of this activity is to read the text on the receipt and to apply the correct metadata to the different parts of the text. The latter means that data must be linked to the correct variables: 'Aldi' is e.g. a shop name and 'Milk' is e.g. a product name.

This activity can be done at two moments in your process:

- On the smart device when the respondent is photographing the receipt or uploads the receipt.
- In the back office where the photos of the receipts or the uploaded e-receipts can be scanned during the 'processing phase'.

  The latter can be applied when:

  - You do not want the result of the OCR/NLP at all to be fed back to the respondent, or
  - You do not want the result of the OCR/NLP to be fed back in real-time to the respondent, or
  - You provide the option of paper diaries with paper receipts.

The activity is supported by the SSI **microservice 'OCR'**.

This activity needs an 'OCR model'.

An option is also that an external party is used to OCR the data. E.g. at SSB, respondents take a photo from a receipt, but this image is externally processed by the data extraction platform Veryfi.

Actor: system

#### 5.2.7.3 Redirect OCR/NLP results to the respondent

The goal of this activity is to show the OCR/NLP results to the respondent. Whether you will actually do this in your NSI, is of course a choice.

And if you decide to redirect a result back to the respondent, you need to decide what exactly will be shown. E.g., when the receipt says 'hlf full mlk' do you show literally 'hlf full mlk' or another predefined label like 'half full-cream milk'?

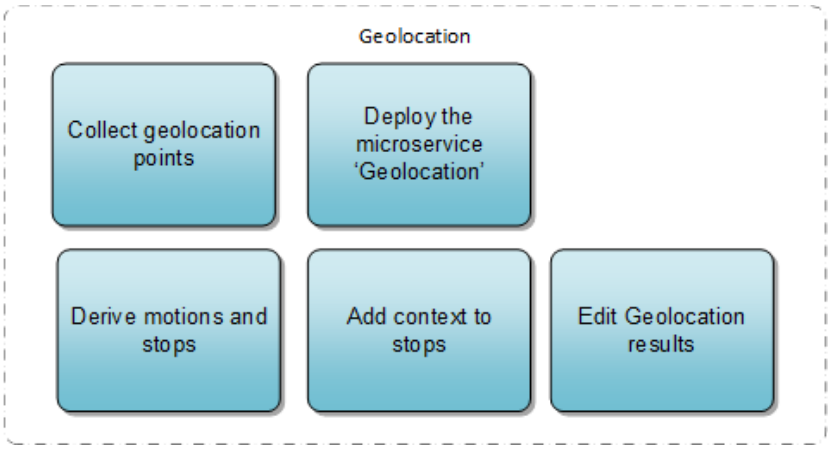
Actor: system

#### 5.2.7.4 Edit OCR/NLP results

The respondent can edit (supplement, correct or delete) the OCR/NLP data, to increase the data quality. Because the output (receipt information) might contain several text mistakes or might have blanks, it might be desirable to let the user correct those mistakes and/or supplement data. Whether or not user editing is desirable also depends on the quality of the input the COICOP classification algorithm needs.

Actor: respondent

### 5.2.8 Geolocation (in case of TUS)




This group of building blocks is relevant when you are going to use geolocation in your process.

Geolocation includes collecting geotracking points, deriving motion and stops, and adding context (e.g. nearby shops).

The premise is that this is done in-survey. The results of the geotracking (motions, stops and context) are namely necessary to support the respondent in recording his activities.

#### 5.2.8.1 Deploy the microservice 'Geolocation'

The microservice 'Geolocation' is provided by the SSI project. As an NSI you are encouraged to use this microservice. This microservices uses also the ML-model 'Geolocation'.

From SSI perspective the advice is to see the microservice as part of the app. In that case the microservice is deployed together with the app in the app store. If the microservice is not part of your app, it needs to be deployed separately, to the platform to be used.

Actor: IT architect, IT developer

#### 5.2.8.2 Collect geolocation points

The goal of this activity is to collect geolocation points. The premise is that the mobile app collects GPS geolocation points. This is not part of the microservice, but it is part of the smart solution in which the internal GPS sensor of the device is used.

*Remark: that is why this solution is an 'internal sensor solution' as mentioned in the smart solution taxonomy.*

Actor: System

#### 5.2.8.3 Derive motions and stops

Based on the collected geolocation points, a machine-learning model (see 'Geolocation model') will derive motions (also mode of transport) based on stops and tracks.

The activity is supported by the SSI **microservice 'Geolocation'**.

Actor: System

#### 5.2.8.4 Add context to stops

In TUS, the respondent must record what activities he performed. To help the respondent, an algorithm could be used to determine (or suggest) the activity the respondent performed. In that case, context could be very helpful. E.g., that a stop was near/at a restaurant.

The activity is supported by the SSI **microservice 'Geolocation**'; this microservice has the functionality to add this kind of context.

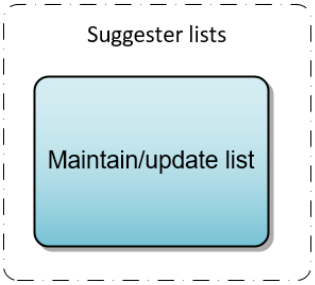
Actor: System

#### 5.2.8.5 Edit Geolocation results

The respondent can edit (supplement, correct or delete) the geolocation data, to increase the data quality. Because the output (motion, stops, transport mode) might contain mistakes or might have blanks, it might be desirable to let the user correct those mistakes, supplement or delete data. Whether or not user editing is desirable also depends on the quality of the input the classification algorithm needs.

Actor: respondent

### 5.2.9 Suggester lists




A suggester list is a collection of items that is used to present (suggest) to respondents, from which the respondent can pick the relevant item. For HBS these are product and shop lists. For TUS these are lists with activities and places. For the Energy Usage use case, this could be a list of appliances.

#### 5.2.9.1 Maintain/update list

The list should be maintained/updated. It should hold all possible relevant items. The list can be used in two process activities to support the respondent:

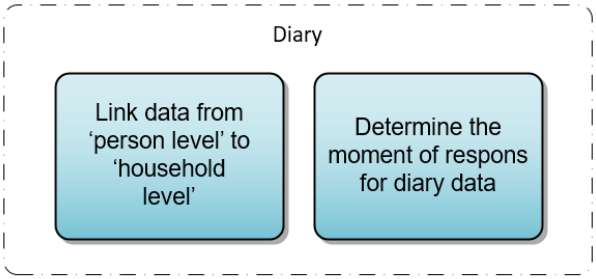
- Correct/supplement results from the smart features (like results of OCR/NLP or location determination)
- Enter information manually by the respondent

It could also support the post-survey process, e.g. for automatically or manually coding.

Concerning a product list: Probably you will choose to link your product list to the COICOP classification.

Actor: Content expert, Data scientist

### 5.2.10 Diary




Activities in this group are all activities to think about regarding the use of a diary, e.g. in HBS and TUS.

#### 5.2.10.1 Link data from 'person level' to 'household level'

In HBS you could implement the diary as one instance for the whole household. In that case all data of all household members is put in one diary. Of course this raises questions about privacy.

On the other hand, you could implement one instance of the diary per household member. In the latter case, the data from each individual person's diary should be linked/joined to the household diary (not for viewing by the household members, but for data processing purposes). The goal of HBS is namely to have the diary data on the household level; there is no need for the NSI to have insight into the expenses per person.

Actor: System

#### 5.2.10.2 Determine the moment of response for diary data

The diary should record the expenditure (HBS), activities (TUS) or the appliances used (Energy Usage use case) of all days in a given period. In principle, in case of HBS, the respondent is free to decide when to record the data (as long as it is within the collection period). Some respondents will do so immediately (i.e., several times a day), others will do it once a day or maybe even once in the whole period. Even after the reporting period ends, the respondent is given a few days to record or adjust expenses.

So, when is the data available for the Process phase? What is the trigger to make the data available?

When using a CAWI questionnaire there is usually a 'send' button. If the respondent uses this, it indicates that they have finished the questionnaire, and the data can be made available for the data processing stage.

For the diary, alternatives can be envisaged:

- Also use a 'send' button. When respondents use this, they indicate that they have finished entering data in the diary. There are drawbacks, however. Maybe respondents forget to push the button. Maybe they have entered data for 5 days (instead of e.g. 7) and did not return to the diary afterwards or they did enter data for all days and just did not use the button.
- Use a button[12] for every day. In this case, the respondents are expected to indicate, for each day in the diary, that they are finished with that day. In this case, the trigger is
    - when all days are 'finished', or
    - make the data available for each day when the specific day is indicated as 'finished'
- Use a time trigger. In this case, the idea is that 'response' is not indicated by a specific action by the respondent, but that it is an automatic time trigger. Specifically, at the end of the

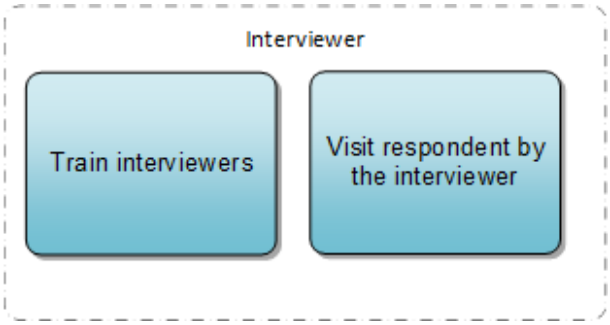
observation period (i.e. x days after the reference period), all data recorded in the diary is considered as response. That is the moment when the data from the diary is going to be made available for the Process phase.

- No explicit trigger: Send the data continuously to the back office, where it is available for processing. However, in this case you need to be aware that data is not yet final/committed. The respondent can still edit the data.

E.g., at CBS for HBS, we choose the 'time trigger'. That gives much flexibility to choose what data to use for different goals. E.g. for the goal of making statistics, use e.g. only respondents that entered data for each day. And for the goal of giving incentives, choose all respondents that entered data for e.g. at least two days. And for analysing purposes, use all registered data of all respondents.

Actor: Methodologist, Survey Designer, Business architect, IT architect

### 5.2.11 Interviewer




In some NSIs, interviewers may play a big role in both recruitment and retainment for smart surveys, but also in the usability of the app, while other NSIs may for various reasons choose not to use interviewers.

#### 5.2.11.1 Train interviewers

You need to get a clear vision of the knowledge that an interviewer should possess. Should the interviewer e.g. support the respondent by downloading and installing the app? Or by using the app? In this case, an interviewer who has more technical knowledge is necessary.

You also need to decide how to train interviewers. Maybe by a (human) trainer, by a manual, by video training, etc.

Actor: Trainer, Interviewer

#### 5.2.11.2 Visit respondent by the interviewer

The goal of this activity is that the interviewer visits the household and tries to motivate the household to participate. Next to motivating, the interviewer can also give instructions on how to download and install the app or even help with installing the app.

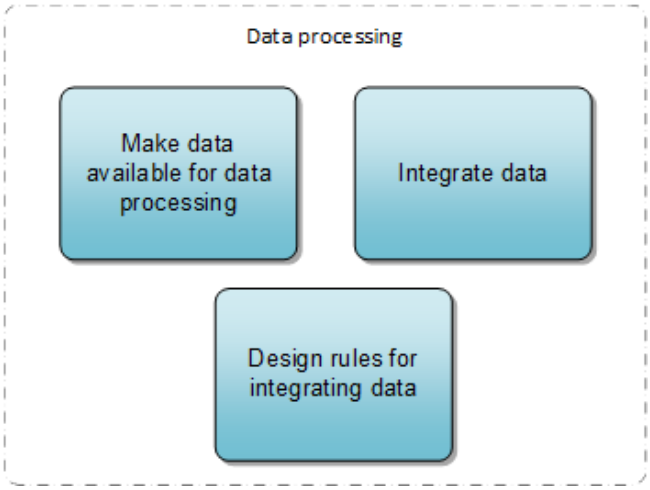
If your NSI also uses paper diaries (as an alternative to the app), the interviewer can e.g. combine this activity with delivering the paper diaries to the household.

Other possible roles of the interviewer:

- Keep contact and motivating the respondent during the survey period.
- Interviewers instruct the respondents over phone at the end of the telephone interview, and after the call, sent credentials for the app and a link to an instruction video by e.g. SMS or email.
- Distribute and help installing the dongle on the smart meter and connect it to the app (Energy Usage use case).

Actor: Interviewer

### 5.2.12 Data processing




This group of building blocks is relevant for the Data Processing (in GSBPM: the Process phase).

#### 5.2.12.1 Make data available for data processing

Data that was recorded in the app during the data collection should be made available to the data processing process. In addition, the survey (e.g. HBS and Energy Usage use case) could also include a questionnaire. So, both sets of data should be made available.

Data should go accompanied with metadata, at least descriptive metadata. So, also for the data collected through the app, there should be metadata available. This may require separate attention.

E.g. at CBS we use Blaise for our questionnaires. The (descriptive) metadata is available in Blaise. However, the app e.g. used in HBS, is not developed with/in Blaise. We had to define an additional activity to create the metadata. Actually, also the microservice used, has its own metadata. This should be integrated with the metadata of e.g. the diary.

Actor: System, Metadata developer

#### 5.2.12.2 Integrate data

In GSBPM the sub-process '5.1 Integrate data' is described as follows: "*This sub-process integrates data from one or more sources. It is where the results of sub processes in the "Collect" phase are combined. The input data can be from a mixture of external or internal sources, and a variety of collection instruments, including extracts of administrative and other non-statistical data sources. […] The result is a set of linked data. […]"*

The question is where both sets of data will be joined together. This is a choice of the NSI. The data could be joined together before the data is made available to the data processing phase, so during the data collection stage. Or it will be a process activity within the data processing stage.

*See also deliverable 2.2 of WP2 where some insights are shared regarding combining smart and traditional survey methods, mode effects and other data integration considerations.*

See also the process building block 'Analyse mode effects' below.

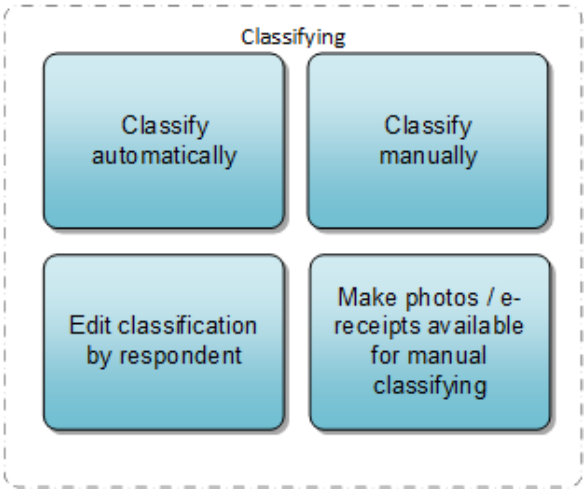
Actor: System

#### 5.12.12.3 Design rules for integrating data

Rules need to be designed to be able to integrate data. For HBS this specifically means that some thinking needs to be done about how to integrate app-data and other data. And also, at what point in the process the 'integrate' activity should be carried out.

Actor: Methodologist, Processing expert

### 5.2.13 Classifying (to COICOP / Activity coding list)




The activities in this group are relevant when classification is needed in your process. For HBS this would be classifying to COICOP. For THUS this would be classifying to ACL ('Activity coding list').

#### 5.2.13.1 Classify automatically

The goal of this activity is to classify each item automatically.

For HBS each article/service from the diary should be classified to a COICOP class. In this case this is done by an automatic algorithm provided by a 'COICOP model'. As input the result of the OCR/NLP activity is used (see activity 'OCR/NLP the receipt' and 'correct/supplement OCR/NLP results').

For TUS each activity from the diary should be classified to an ACL class. In this case this is done by an automatic algorithm provided by an 'ACL model'.

*Remark: also, string-matching could be used. We did not describe that yet in this document; we have to look into that for a next version of this document.*

As an NSI you need to decide whether you want the result of this classification to be fed back to the respondent (see also the activity 'Decide on app architecture'). In that case this activity needs to be done somewhere in the data collection phase. It requires namely a short feedback loop, and the data is not yet final/committed.

If you do not want to lead the result of the classification back to the respondent, then this activity can be done in the data processing phase.

An NSI could also decide to do both.

The activity is supported by the SSI **microservices 'COICOP classification'**[16] and **'HETUS classification'**[17]

Actor: System

[16] This is part – together with the 'OCR' microservice – of the 'Receipt scanning' microservice

[17] This is part – together with the 'Geolocation' microservice – of the 'GeoService' microservice

#### 5.2.13.2 Classify manually

In this case, classifying is not done automatically, but manual by a coder. Normally this activity is done next to the automatic classifying. Probably only items (articles or activities) that cannot be classified automatically are classified manually.

In case HBS, as input the result of the OCR/NLP activity is used and/or the actual photos/e-receipts (see 'Make photos / e-receipts available for manual classifying').

Actor: Coder

#### 5.2.12.3 Make photos / e-receipts available for manual classifying

If your NSI decides to use photos of receipts and/or e-receipts to manually classify to COICOP, an activity is needed to provide them. This means that photos/e-receipts should be stored somewhere so that they are accessible for the 'Classify manually' activity.

Actor: System

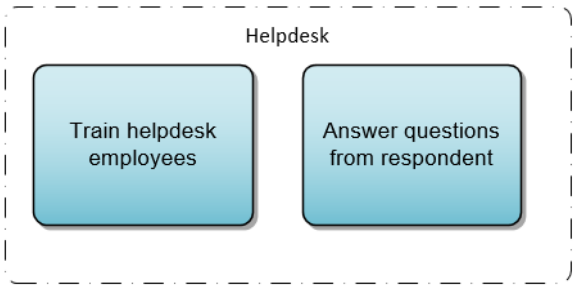
#### 5.2.12.4 Edit classification by respondent

The respondent can edit (supplement/correct/delete) the result of the classification, to increase the data quality.

This means that the classification (the list with classification codes) must be available in or by the app.

Actor: Respondent

### 5.2.14 Helpdesk




The activities in this group are necessary if you want your helpdesk to be able to support respondents in using the smart solution. With regard to smart solutions the process activities in itself are not different than for other modes (e.g. you still have to handle incoming emails or phone calls), but the knowledge needed to respond to the questions is different.

We could distinguish between three types of helpdesks: Substantive (about HBS/TUS itself, why and how), about the app (installing, using) and technical (bugs, issues, errors). For the substantive helpdesk there is nothing different in comparison with the non-smart version of the survey. For the other two types, the activities below need attention.

#### 5.2.14.1 Train helpdesk employees

This activity does not need much explanation. You need to get a clear vision of the knowledge that a helpdesk employee should possess. Should the helpdesk e.g. support the respondent by downloading and installing the app? Or by using the app? This requires specific capabilities of a helpdesk employee.

And then you need to decide how to train them. Maybe by a (human) trainer, by a manual, by video training, etc.

Actor: Trainer, Helpdesk employee

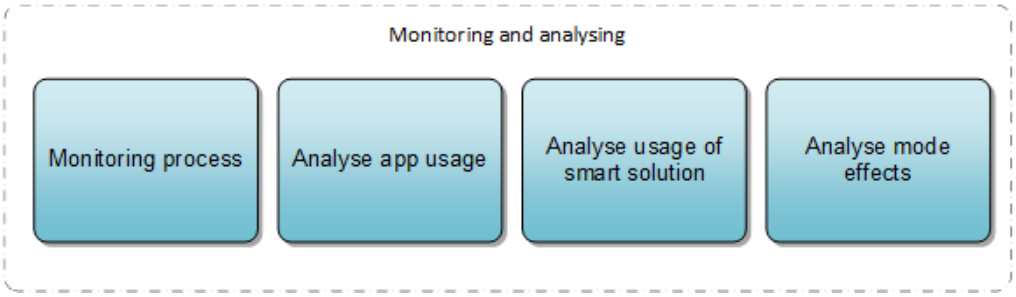
#### 5.2.14.2 Answer questions from respondent

The goal of this activity is that a helpdesk employee supports the respondent by doing certain tasks. From the ‘smart solution’ aspect the focus is mainly on the usage of the app or other type of smart solution.

Depending on what type of support is necessary, this activity gets a certain place in your business process.

Actor: Helpdesk employee

### 5.2.15 Monitoring and analysing




The difference between monitoring and analysing is that monitoring is aimed at adjusting the ongoing production process (collection cycle), e.g. to send an extra reminder when response rates are disappointing. Analysing is part of the bigger learning cycle to improve the methods, processes, tools for the next run of your production process.

#### 5.2.15.1 Monitoring process

The goal is to monitor the ongoing survey, with regard to the process. Paradata is necessary as input.

> *"In general, a smart data quality monitoring system should be aimed at controlling the data collection process, in order to intervene in real time when problems impacting on the survey progress and on representation and measurement errors occur." Source: [3] §3.1.3.2. "Smart data monitoring and use of contextual data".*
>
> *In [3] §3.1.3.2. a monitoring system for sensor data is described considering indicators to monitor survey progress and a paradata system that should be developed at micro-level. Some of the indicators are shown below as an example.*

Monitoring indicators are e.g.:

- Number of respondents downloading and/or installing the app
- Number of participants who drop out of the survey
- Number of calls to the helpdesk concerning the smart solutions
- Number of responses (response rate)
- The (average) number of days for which data is provided (so far) in the diary by a respondent
- The (average) number of receipts registered (so far) in the diary by a respondent

Actor: Monitoring officer

#### 5.2.15.2 Analyse app usage

The goal is to get insights into how the app is used by respondents, to learn from it and use it in the PDCA-cycle.

Indicators are e.g.:

- How often is the app started/opened (usage frequency)
- How do respondents navigate through the app
- At what point do respondents close the app
- How often do respondents get technical errors
- Number of respondents de-installing the app during the survey
- Android versus iOS usage
- Type of device that is used

The question of course is whether this is technical feasible.

And also here, the privacy questions arise. Is it allowed to collect all these data? What level of detail is allowed to collect this kind of data? Who in your NSI may have access to this data? How long may you keep the data? You need to think of a process to delete/destroy the data.

Actor: Analysing officer, Privacy officer

#### 5.2.15.3 Analyse usage of smart solution

The goal is to get insight into the usage of the smart solution itself and the quality of the registered data.

E.g. in case of HBS:

- The distribution of receipts registered by using photos, uploading or using manual entry
- The same for the distribution per registered product line
- The number of times the respondent manually updated the OCR/NLP result
- The number of times the respondent had to retake a photo to get a good-quality scan

Of course, it would give extra insight when figures are distinguished by background variables like age, education, etc.

But of course, also here, the privacy questions arise and need to be addressed.

Actor: Analysing officer

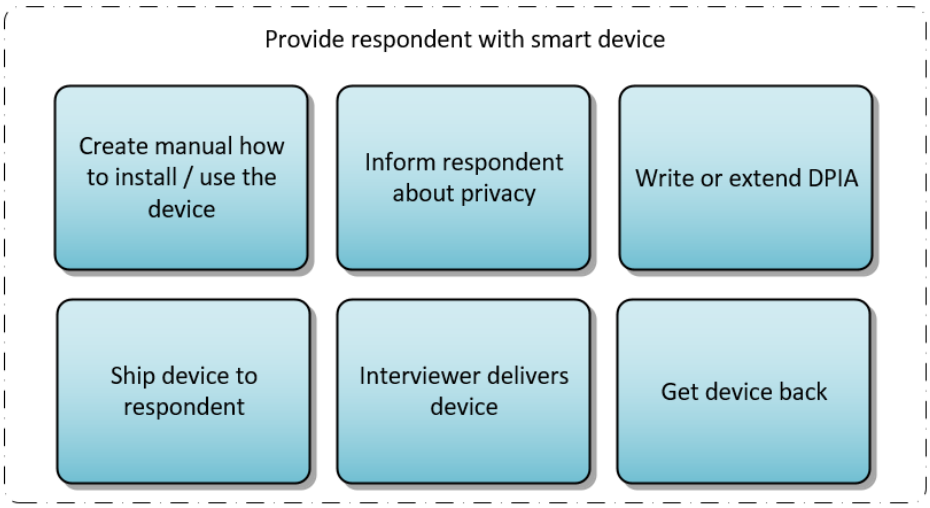
#### 5.2.15.4 Analyse mode effects

Using an app for data collection is seen as a new mode. This gives reason to pay extra attention to mode effects.

Reference [3] states *"Previous research has demonstrated the utility of smart surveys as a methodology for addressing the limitations of traditional survey methodology by substantially reducing respondent burden and enriching the data. The most significant obstacle to this natural progression arises from the concerns of how the new data streams from the smart surveys can be integrated with the old. Even subtle differences between modes of data collection can potentially lead to differential measurement, and the differences in data format that may arise from passive data collection or sensor measurements may not be easily reconcilable. To this end, next steps in this area must involve the intentional estimation of mode effects to establish where the vulnerabilities are greatest. Depending on the degree and location of differences, various integration techniques may be appropriate, ranging from direct integration to complex methodologies."*

Actor: Analysing officer (Methodologist)

### 5.2.16 Provide respondent with smart device




These are process activities that are applicable when you want the respondent to use a smart device that is not already in possession of the respondent. E.g. an energy dongle or an activity tracker.

In the (pilot) use case of SSI a dongle and a third-party app were used. The energy data usage was visible through the app. To make the data available for the NSI, the respondent had to download the data through the app and send it by mail to the NSI. Respondents were asked to keep a (CAWI) diary throughout this period, recording 1) Their presence at home (e.g., which parts of the day they were away or at home) and 2) Which major electrical appliances were used (e.g., dishwasher, washing machine, tumble dryer).

#### 5.2.16.1 Create manual on how to install / use the device

In this case the respondent is unfamiliar with the device. This activity is to create a manual for the respondent on how to install and use the device. This is very essential. If the respondent cannot install the device, responses get lost. And not using the device properly will result in no or poor quality of the data. So, it is essential that communication experts are involved. Technical knowledge of the device is essential as well.

Actor: communication expert, technical expert

#### 5.2.16.2 Inform respondent about privacy

In this case the respondent will have more privacy concerns, because the respondent is not familiar with the device. Additional attention is necessary, to address the risk of losing responses. The NSI has to think about how to inform the respondent about the privacy aspects.

In the SSI use case of energy usage special attention was paid to Data Sensitivity & Informed Consent. "*Energy usage data can reveal detailed patterns about household occupancy and routines. Therefore, all participants were fully informed about the purpose of the study and the types of data collected. They provided consent to share their energy usage data and diary entries. The potential privacy implications were explained, and participants were given the option to withdraw from the study at any time.*" [source: End report of 'SSI Energy use case: Dongle pilot']

Actor: communication expert, legal officer

#### 5.2.16.3 Write or extend DPIA

In this case the DPIA should be extended (or created if not present). There are additional privacy and security risks when a respondent should use a device provided by the NSI.

In this case you could use the modular strategy that has been described by WP5 for developing the DPIA (see reference [7]).

Actor: legal officer, security officer

#### 5.2.16.4 Ship device to respondent

In this case the device (including the manual) is shipped to the respondent. The device should be packed and shipped. Also consider the situation where the package could not be delivered. The question is also whether the NSI does the packaging and shipping itself or that a third party is involved.

Actor: logistical staff

#### 5.2.16.5 Interviewer delivers device

In this case an interviewer delivers the device to the respondent. That has the advantage that the interviewer could help the respondent with the installation and initial usage of the device. The interviewer needs additional knowledge (see activity 'Train interviewers').

In this case, one also needs to think how the devices (and the manuals) get to the interviewer.

Actor: interviewer, logistical staff

#### 5.2.16.6 Get device back

The question is whether, after the survey, the NSI would like to get the device back. In that case the NSI needs to arrange things so that the respondent is willing and capable of returning the device to the NSI. Maybe you need to provide a return box to the respondent. Also, to increase the return rate, the respondent should not have to bear any shipping costs.

Another choice is that the interviewer goes to collect the device.

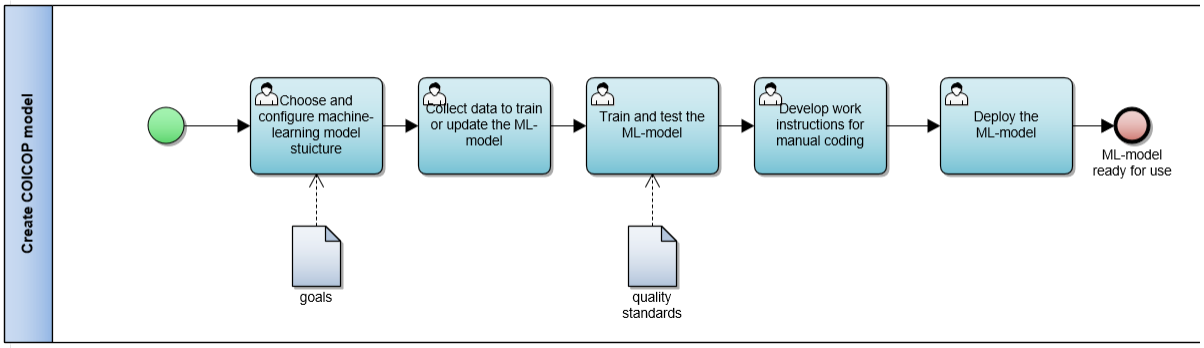
Actor: interviewer, logistical staff, financial staff

## 5.3 Business process example

This is an example – in the context of HBS – of a business process model. The model is built with the help of the building blocks. The example is limited to activities that are affected with respect to smart solutions. E.g., an activity 'Design invitation letter' is not modelled. Probably the content of the letter will hold something specific when using an app, but the activity in itself is not different.

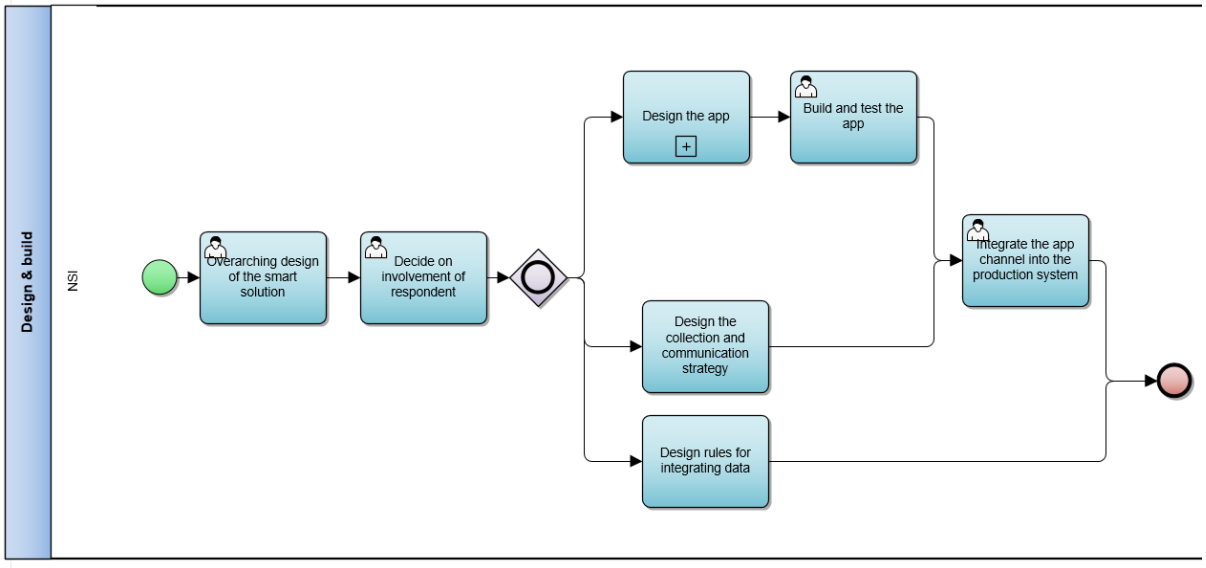
To be clear, this a fictitious example. It is not an example of any NSI. The only purpose of this example is to show how you can use the building blocks. Therefor it is also not a model of a complete business process; we just modelled some parts which are sufficient for the purpose.

**Create COICOP ML-model**



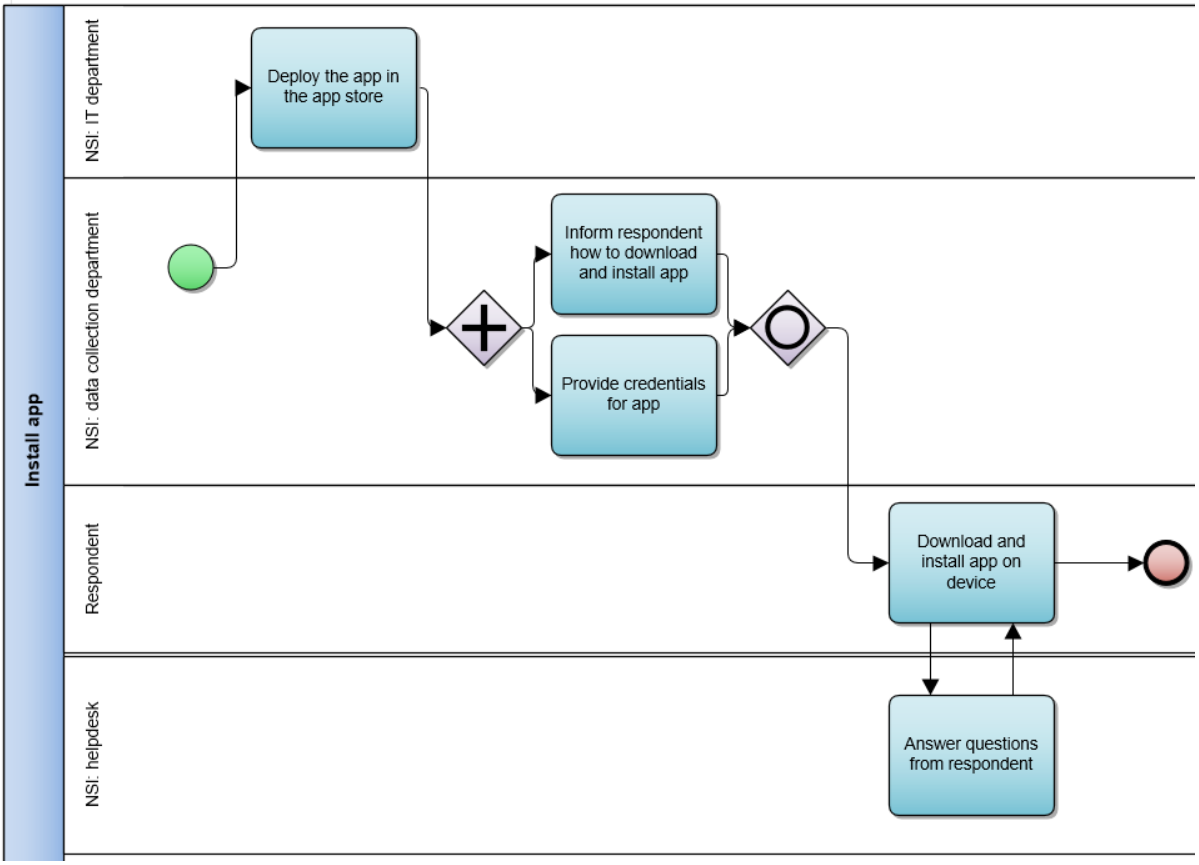

This is the 'pipeline' to create the COICOP ML-model that will be used in the production process. As we will see below, in this example-process, the classification is done post-survey.

**Design & build**




Based on the needs, the major decision is how the overarching design of the smart solution will look like. Closely related to this is the decision on how to involve the respondent in the 'respondent journey'. Based on that the different design elements will be designed. The design of the app actually exists of three activities (see the building blocks); for simplicity they are not mentioned separately in the model. Based on the design, the app is built and tested. Also, the collection strategy is designed and the communication strategy. In the end the app channel is integrated into the production system.

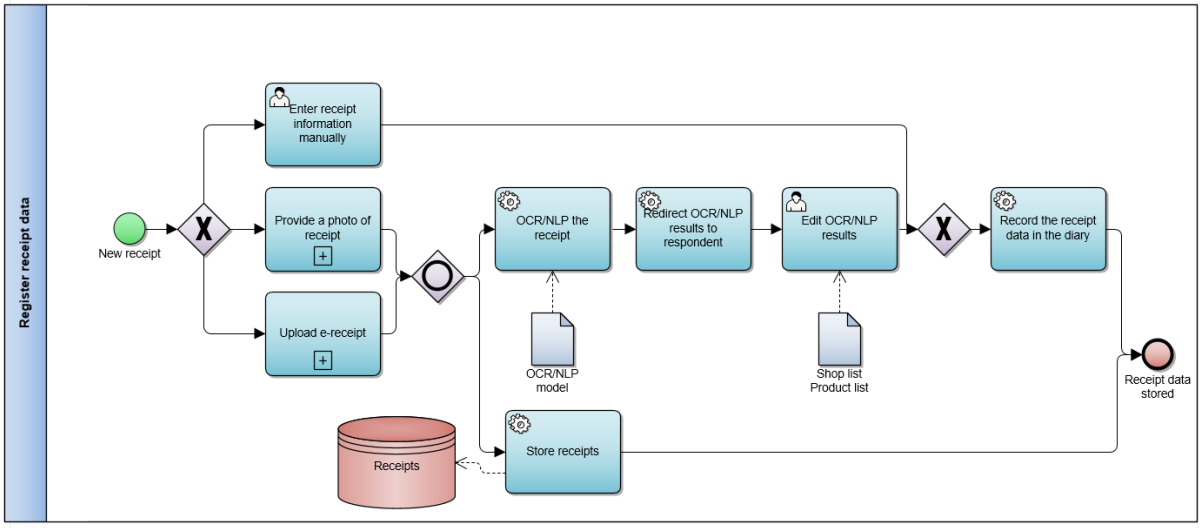
**Install app**




In this example, the data collection department triggers the process. They tell the IT department to put the app in the app store. When that is done, the data collection department informs the respondent how to download and install the app, and they will also provide the credentials for the app to the respondent. This could be done by putting all this information in the invitation letter.

Then the respondent can download and install the app and use it for the first time. For help, the respondent can contact the helpdesk. So, in this example, the interviewer has no role in helping the respondent.

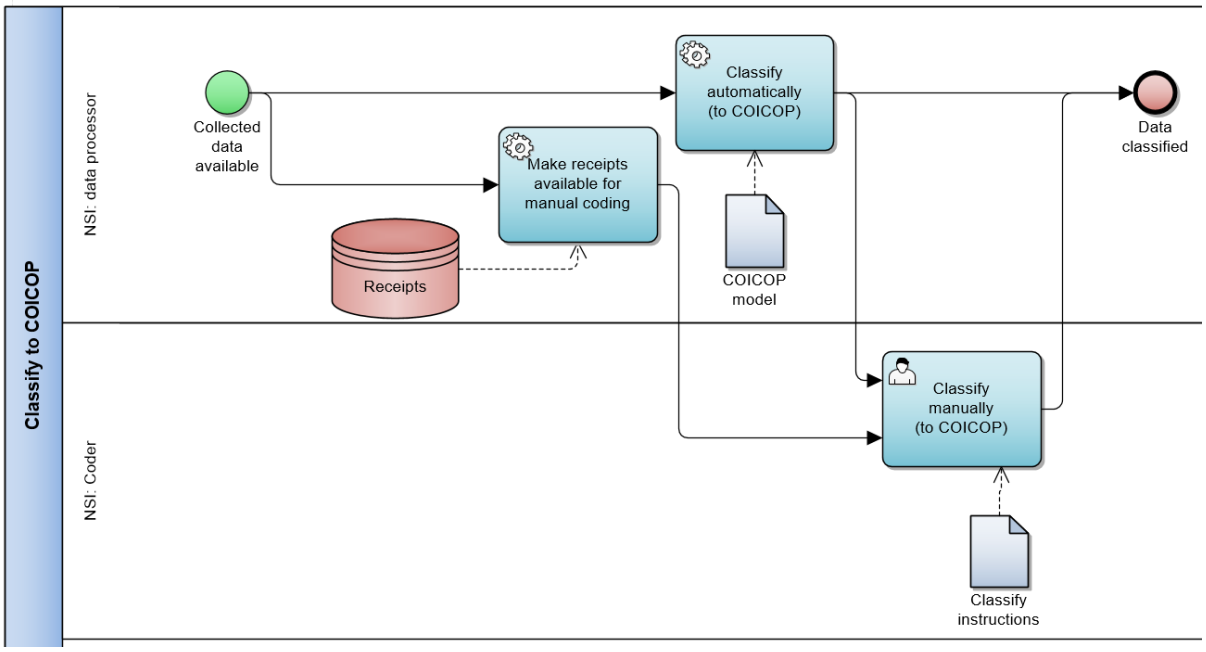
**Register receipt data**




In this example has three ways of providing receipt data: manually, through a photo or through an upload of an e-ticket. Photos and e-tickets are stored. Each receipt is going through the process of OCR/NLP. Apparently in this example, the OCR/NLP results are fed back to the respondent. The respondent can then edit the results. All the information is then stored in the diary.

Note that in this example there is no in-survey classification to COICOP.

**Classify to COICOP**




In this example process, the classification is done post-survey. The first is to try to classify as much as possible automatically. For this, the COICOP-model is used (that was created in the 'Create COICOP ML-model' subprocess, see above).

Receipts that cannot be classified automatically are then classified manually. For this, in this example, the coder has access to the receipts (photos and e-tickets).

# 6. PDCA-cycles for smart surveys

## 6.1 Introduction

Part of the activities of a NSI is to perform governance. This enables the NSI to deliver the statistics needed within regulations and in an environment that is changing continuously. Part of the environment of the NSI are the customers. They are in demand of the various surveys the NSI is entrusted to deliver.

In order to be able to deliver the demanded statistics, the NSI continuously scrutinizes the applied way of working and the surroundings of the organisation. Thus, the NSI is able to determine changes necessary to stay aligned with its environment. In this manner, the NSI safeguards that targets are met and that delivery is performed within the boundaries set. Necessary changes could cause alterations to the relevant aspects of a survey or to the processes or to the organisation of the NSI. The governance activities of a NSI, are part of the “Strategy and Leadership” activities as determined in the GAMSO model.

Within this work package, we consider IT-architecture, methodology, logistics and legal-ethical as relevant aspects that are critical to the ability in developing and using a smart solution. By incorporating these aspects in the development and use, we assure that a smart solution is fit to be used in a smart survey.

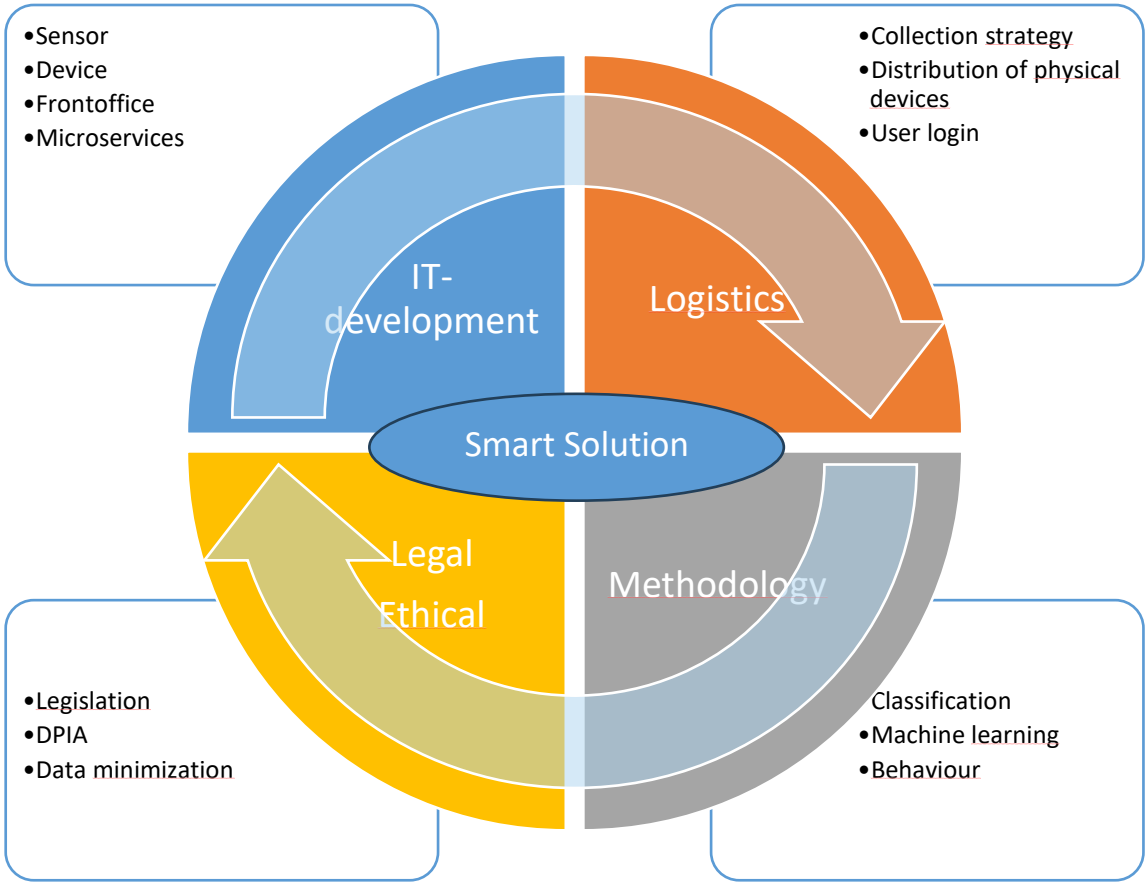
As mentioned above, each NSI has an existing, overall construct that enables it to be in control and perform operations within regulations and targets set. Typically, this construct uses PDCA-cycles on various levels within the NSI and for the various aspects relevant. During a survey, different components are used in a specific constellation and with specific parameters configured. All these components have their individual PDCA-cycle. When developing and implementing a smart solution, new PDCA-cycles need to be implemented for this new solution.

For each survey performed, the NSI evaluates whether the used collection methods and applied methodologies, still are fit to create statistic at hand. This evaluation is part of a the overall general PDCA-construct of the NSI.

Within the GSBPM model, the PDCA-activities are performed within the overarching phases of “Specify needs” (Act), “Design and Build” (Plan), “Collect, Process, Analyse and Disseminate” (Do) and “Evaluate” (Check).

## 6.2 Development of a smart survey

When a new smart solution is to be developed, it will change the current abilities of the NSI. The four aspects will affect each other and the final smart solution. Requirements desired by one aspect could be blocked by the requirements needed by another aspect. For example, the desire for an on-screen explanation on a smart device could be blocked by the physical dimensions of the screen of the device.

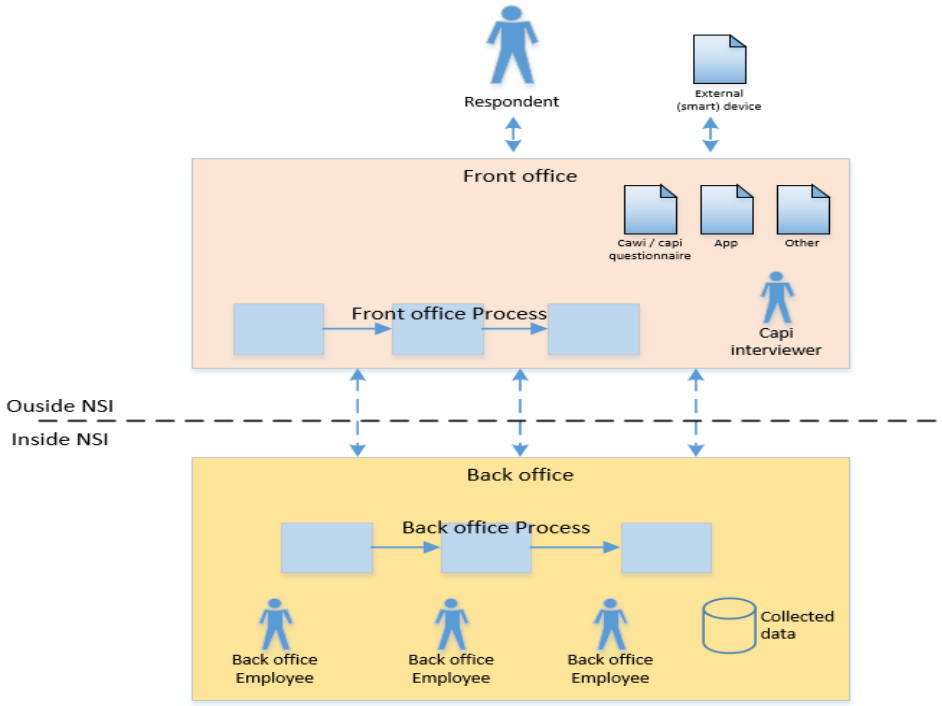



Because of the complex interaction between the various aspects, CBS (as example) has established a dedicated working group during the initial development and implementation of a smart solution. This group assesses, coordinates, and guides the process in an iterative manner from multiple aspects, aligning them with the various process steps of the GAMSO and the GSBPM framework.

In the figure above for each aspect of the development and use of a smart solution, some topics are mentioned that affect the development of the smart solution. In the next paragraphs we will elaborate on the PDCA-cycle for topics mentioned in the logistics and the methodology aspect (the latter focuses on machine-learning models). In these descriptions, we will explicitly mention the assumptions made and the current ways of work applied, that need to be validated during the PDCA-cycle. This evaluation will help us determine whether assumptions and ways of work enable us to reach the desired outcome within our boundaries. For a PDCA-cycle within Legal-Ethical we refer to the WP5 deliverable (see reference [7]).

This structured PDCA approach helps ensure the smart solution aligns with NSI standards, meets user needs, and remains adaptable for future advancements.

## 6.3 PDCA and Logistics




The introduction of new smart solutions will be accompanied by a redesign of existing approaches within the logistics process. After all, the smart solution introduces new methods for contacting respondents and collecting data thus a new collection strategy will come to life.

Within the PDCA cycle of the logistics process, the development of the survey that is using the new smart solution goes hand in hand with the development of the collection strategy. This is to ensure that both internal and external requirements are considered. Here, "requirements" should be understood as what is logic to the respondent in relation to the purpose for which the smart solution is used, as well as whether it meets the various internal requirements set from different perspectives by the different interested parties.

### 6.3.1 Plan Logistics

During this stage, the various needs and objectives that define the purpose of using the specific smart solution are identified. The stakeholders and interested parties are engaged and their requirements are gathered. Next to this the availability, completeness and accessibility of the data is planned. All these requirements lead to an initial design of the collection strategy and this design is assessed on its feasibility.

Within the current smart surveys developed, like HBS, switching between modes has not been included in the design. In the initial versions, the main focus of the research is the acceptance rate of respondents and the new structuring of CBS capabilities. In a later iteration of the PDCA, impact of the enabling switching between devices or modes, could be part of the research conducted. In this case our assumptions and way of work on the use of multi modes within a smart survey can be researched.

The use of sensors enables us to have a continuous stream of data provided by the respondent. In the logistics process, the decision has to be made at what moment gathered data will be used for further processing and whether a respondent gets to verify the correctness. In the HBS survey, the data that has been extracted from the photo of a receipt is not classified by the respondent. This way of work is based on research done by other NSI's which shows that respondents do not expect that they need to verify whether the photo of a receipt is translated correctly to classified products.

Within non-smart surveys, only data is processed when a respondent explicitly has granted its permission to use it. Mostly this explicit permission is granted by using a "send" button. The same applies to the HBS. Only data on expenditures is processed when the respondent explicitly has granted its permission by submitting its diary on expenditures to the NSI. In TUS, the data provided by the sensor is used to support the respondent to verify the nature of his travels and the possible destinations. In this case data of the smart solution is only used after explicitly being validated by the respondent.

### 6.3.2 Do Logistics

In this stage, the plans are brought into action. Here we use the concepts of a minimum viable product to carry out one or more pilots. The solution is introduced to a small-scale test environment to gather feedback from internal users and pilot participants. In the end the solution is deployed into production and constantly monitored. For this monitoring the various key performance indicators are designed and implemented.

New to the survey on energy consumption, is the use of a dongle. The device is connected to a smart energy meter and is able to provide various readouts on the use of electricity and natural gas. Although CBS has experience in providing physical devices to be used by a respondent, in this case CBS also has to provide some sort of support for connecting the dongle to the smart energy meter and installing the dongle onto the Wi-Fi-network of the respondent. CBS has decided at first not to test the process of ordering and distribution of such a device on a large scale. Only a few volunteers from CBS are provided with the dongle. However, they will use the manual on connecting and installation in order to provide feedback.

### 6.3.3 Check Logistics

In the Check stage the actions taken and outcome produced during the do stage are evaluated. This is in order to determine whether they align with the objectives and expectations. The various performance metrics on accuracy, efficiency and usability of the smart solution are assessed and the data is validated. This ensures that the statistics produced are up to the standards of the NSI. Next the feedback on and adoption of the new survey is evaluated. And the business process in place is scrutinized.

As mentioned before, based on research done by other NSIs, respondents do not expect that they need to verify whether the photo of a receipt has been translated correctly into the products bought. The assumption that they would find this normal was proved wrong during various observations when researching the behavior of the respondents using the solution on their devices. This provides an important message when innovating. One needs to be open to check on unexpected outcomes as they reveal some implicit assumptions that could not be applicable.

### 6.3.4 Act Logistics

Based on the various identified areas of improvement, the use of the smart solution is assessed on its merits to the survey. When these merits are positive or are expected to be positive, the solution

is refined by improving algorithms, interfaces with the users and processes. Incorporating these alliterations, the solution is implemented for full-scale use by the NSI.

Finally, during this stage, insights from this survey and other surveys are used to plan the next PDCA cycle. By doing so, one is able to optimize the use of development services, processes and components for more than one survey. On option that could be a change in the future is the use single sign on for various different modes within a survey of even for different surveys.

## 6.4 PDCA and machine-learning models

### 6.4.1 Introduction machine-learning models

Smart Surveys aim to reduce response burden by integrating sensor data or assisting the user in completing the survey smartly. In both scenarios, machine learning models are often used to process the data coming from sensors and predict aspects of the respondent's behaviour. In this sense, machine learning models can improve survey efficiency, improve response burden, or help improve the quality of the response. For example, in the household budget survey, machine learning is used to help reduce the response burden by helping the respondent keep a spending diary more efficiently. The machine-learning model processes pictures that a respondent took from her receipts. First, the scanned receipt is pre-processed, its text extracted using an OCR machine-learning model, and then the function of each text on the receipt is classified using document understanding (product text, shop text, price text, etc.). Second, the result of the OCR model is used to classify the product rows into COICOP product categories using a second microservice. Likewise, in a time-use survey, the GPS sensor of a mobile phone can be used. A machine learning model can then try to classify the respondent's activities based on the GPS location data it receives. The model is used to identify stop/track clusters, etc.

Before a machine-learning model can perform tasks like OCR, classifying receipt texts, or labelling activities, a machine-learning model has to learn to perform these tasks. We call this training a machine learning model. Roughly, there are two ways to train a machine-learning model. On the one hand, we can train a model in a *supervised* manner; supervised training means we provide a model with example inputs and the labels for each input. For instance, to train a COICOP classification model, we need example receipt texts from stores and their corresponding COICOP labels.

On the other hand, we can train a machine-learning model using *unsupervised learning*. Unsupervised learning means we do not have any labels available, and the model needs to look for patterns in the data itself. An example of unsupervised learning is clustering. Clustering models use similarity metrics to discover groups of data points that have similarities. It may have become clear that training models in an unsupervised way, is more difficult than training models in a supervised way. We do not have a "ground-truth" label that we can use to check the model's quality. Both supervised and unsupervised machine-learning models need large amounts of data to train them. In the case of supervised machine-learning models, we also need large quantities of labels, which often need to be created by manually going through the data and deciding which examples have which label.

After training a machine-learning model we hope the model can generalize to previously unseen data; we hope it can discover patterns that help it to classify these unseen examples correctly. More importantly, we hope it keeps classifying unseen examples correctly. The data a machine-learning model sees in a real-world scenario often changes (and keeps changing) from the data it has seen

during training. Stores introduce and remove products from their product inventory every month, which introduces new receipt texts that weren't part of the original training set. In addition, at the time of model development and training, not all data might be available yet. In the case of the HBS app, for example, we may not have data available from all stores. These challenges mean that training a model once will rarely be sufficient. Therefore, the machine-learning models that we deploy in smart surveys, will have to be constantly monitored and updated. In this chapter, we lay out a PDCA cycle that will assist in keeping machine-learning models up to date.


### 6.4.2 The Machine Learning Process

Figure 1 shows the process of training a machine-learning model along with its PDCA stages. It can be seen that there are five stages in the machine learning process: data collection, data cleaning, exploratory data analysis, model building, and model deployment. Each of these stages, maps to a certain stage in the PDCA cycle. Note that not all stages in the Machine Learning Process map completely to the stages in a PDCA cycle. For example, in Figure 1 an exploratory data analysis is part of the "Do" in the PDCA cycle. However, in the "Plan" stage, we would like to define the model performance metrics that will be used in the later stages, but these model performance metrics depend on the result of the exploratory data analysis. In the next section, we will discuss the four stages in the PDCA cycle, along with the stages of the machine learning process that belong to them.

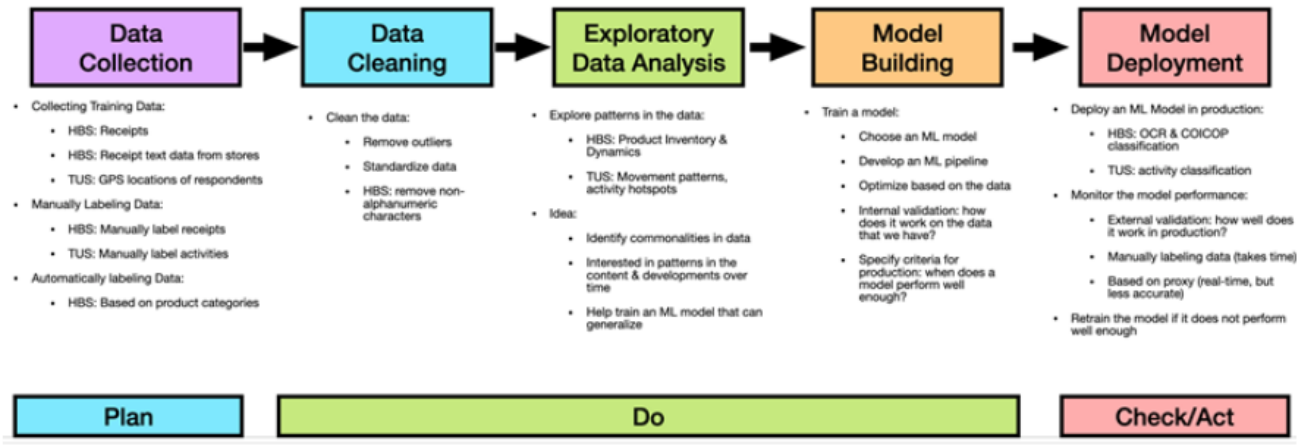



Figure 1: The Machine Learning Process

#### 6.4.2.1 Plan

In the Plan stage we first *define the objectives* of the machine-learning model as part of the smart survey. On the one hand, we specify the goal of the classifier. For instance, we want to accurately categorize receipt texts into their corresponding COICOP classes, or we want to correctly classify respondents' GPS locations into activities. Part of defining the objectives is also specifying what success looks like. One way of specifying success is to choose a performance metric for the machine-learning model and the level of performance that we would like to achieve.

Second, we also define the scope of the project. We may decide that, based on the data we received, or the difficulty of the task ahead, to limit the machine-learning model to specific situations. We may, for example, decide that we limit COICOP classification of receipt texts to supermarkets only, based on the availability of the data, the frequency of the scanned receipts, and the response burden for this specific category of receipts. Likewise, for TUS, we may focus on a limited number of activities relevant to the survey. In addition, when defining the scope, we may decide on the maturity level of the implemented model. Are we developing a model for proof-of-concept, a pilot study, or a full-scale implementation?

Third, the "Plan" stage entails the data collection stage of the machine learning process. During data collection, it is important to identify the sources of data to train the machine learning model. For HBS, for the OCR microservice we need to collect receipts from various stores. For the COICOP classification microservice, we can collect receipts from the various stores, or we can try to acquire the receipt texts from the stores directly. Likewise, for TUS, we may ask a select group of people for permission to monitor their activities during a certain period of time. During data collection, it is also important to know whether the model is an unsupervised or a supervised machine-learning model. A supervised model needs labelled data, and this means that we need to make sure these labels are acquired. Sometimes, labelling can be done automatically using existing data sources. In the HBS survey, for example, we can use CPI scanner data to link COICOP labels to receipt texts. However, often manual labelling is necessary. This involves human annotators going through the data and manually specifying the label for each training example. As this is a very labour-intensive activity, sufficient time and resources must be allocated for this activity. During data collection, a first impression of the quality of data can also be obtained. Having said that, a more detailed analysis of the data quality will be carried out during the data cleaning and data exploration phases, as part of the Do stage.

Fourth, already a preliminary choice of *machine-learning algorithms* can be made. Based on the data collected, a choice needs to be made between using supervised or unsupervised models. In addition, already a class of models can be chosen as well. For supervised models, a choice can be made between classification or regression models. For unsupervised models, a choice can be made for the type of clustering model. It may also already be possible to investigate further extraction methods and models specific to a task, like text classification or detecting objects in images. The choice for a specific model is made later during model training and evaluation and also depends largely on the data collected.

Fifth, in this stage we *determine evaluation metrics* for the machine-learning model. On the one hand, these are evaluation metrics that measure the model performance during training. For instance, if the receipt texts of a certain category must be classified correctly, even if we do not find all of them, we may choose the *precision* performance metric. If it is more important to find instances belonging to a certain category even if we label some of them incorrect, we should choose *recall* metric. Conversely, the *F1* metric is the harmonic mean of both previous metrics and therefore

balances both of them. The definition of success may vary for each survey, leading to different considerations for selecting relevant performance metrics.

On the other hand, a decision should be made about how the model will be monitored in production. A model in production can be monitored using the same performance metrics used during training or using proxy metrics. Using the same performance metrics used during training means that we need to label new data coming in from production. This means we need to allocate personnel to perform the labelling. Because labelling takes time, using these performance metrics will mean that we cannot monitor the model in real time. If real-time observation is important, instead proxy metrics can be used that monitor data and prediction drift (input and output drift) using statistical methods, distance-based methods, or rule-based checks[1].

Last, based on the results of the previous steps we can *develop a plan*. The steps for preprocessing the data, training the classifier, and evaluating its performance can be outlined. This plan is the input for the following phase: Do.

#### *6.4.2.2* Do

In the "Do" phase, the plans made in the "Plan" phase are brought into action. This phase entails the data cleaning, the exploratory data analysis, and the model-building phases of the machine learning process. We will discuss the "Do" phase in terms of these three phases of the machine learning process.

First, during *data cleaning*, the data is cleaned by removing noise, irrelevant information, inconsistencies, and outliers. Moreover, data from different sources may have different file formats and a different shape or different column names. In this step, the data should also be converted into a standardized format for further processing in the follow-up steps. For the HBS data, this may involve cleaning the receipt texts, removing non-alphanumeric characters, and converting all receipt texts to lowercase. Also, the COICOP labels are added to the data in this step. In addition, some machine-learning models may need specific data-cleaning steps, like stemming, lemmatization, stop word removal, or tokenization (for text data).

Second, in the *exploratory data analysis*, the data is analysed to discover patterns, commonalities, and time-dependent patterns in the data that can be helpful during model training. The exploratory data analysis will give an impression of the ability of the machine-learning model to generalize over the data. For HBS we would analyse aspects like product inventory and dynamics in this step. Comparing the product inventory of different stores gives us an idea of how similar the receipt texts are. Looking at the product dynamics gives us an impression of how fast our model will be outdated. For TUS, the data exploration could look into finding movement patterns and activity hotspots.

Last, during *model training*, a model pipeline is developed, trained, evaluated, and deployed. A model pipeline consists of data preprocessing steps, features extraction methods, and a machine-learning model. During this step, several preprocessing steps, feature extraction methods, and machine-learning models are trained and evaluated. The best-performing models, according to the criteria we defined in the "Plan" phase are selected and deployed in production. In addition, to the criteria defined in the "Plan" phase, we may identify some additional criteria in this step that are relevant to a model in production.

Note that the evaluation of the model is an evaluation of the internal validity. The model is evaluated on a sample of the data that was collected during the data collection. While this gives us

an idea of how good a model generalizes to unseen data, it does not give us a complete picture of its performance in real-world scenarios. Therefore, after initial evaluation, the model must undergo further testing in real-world conditions to assess its external validity (during the “Check” phase).

#### *6.4.2.3 Check/Act*

In the Check/Act phases, the model is deployed into production and constantly monitored. The “Check” and “Act” phases map to the model deployment phase of the Machine Learning Process.

In the “Check” phase the actions taken during the “Do” phase are evaluated to determine whether they align with the project objectives and expectations. In terms of a machine-learning model, this means the deployed model is constantly monitored to detect issues and anomalies in the model running with production data. On the one hand, this involves monitoring the input side of the model, or the data. Monitoring the input side of a machine-learning model involves various checks and analyses to ensure the quality of the data being fed into the system. For example, this could entail examining data distributions over time to detect any shifts or anomalies that might indicate changes in the underlying patterns or characteristics of the input data. Additionally, data validation techniques can be employed to identify outliers, missing values, or other inconsistencies that could potentially impact the model's performance.

On the other hand, this involves monitoring the model output, which involves evaluating the model's predictions and assessing their accuracy, reliability, and consistency in real-world scenarios. The model output can be monitored using model performance metrics like recall, precision, and F1, as we discussed in the “Plan” section. The downside of this approach is that you need labelled data to evaluate these metrics. This means new data available in production needs to be labelled. Labelling is a time-consuming and labour-intensive process and therefore it will not be possible to evaluate a model in real-time with these model performance metrics. In contrast, it would be possible to use some proxy measures that reflect model quality. A change in the distribution of the predicted COICOP classes may indicate something is happening with the model. Then again, season-specific products may influence the label distribution similarly, and model performance may just be fine. Last, it would be possible to get feedback from the users themselves; by enabling them to correct the model predictions we have an additional metric of the model's performance.

Finally, based on the results of the “Check” phase, action may need to be taken to improve the machine-learning model. This may mean that we need to collect additional training data, for instance, for a store that is missing or for a COICOP category that is underperforming. In the “Act” phase we then take the actions to improve the machine learning model, retrain it, and deploy a newer version. For completely new data (a new store or COICOP category), this means starting from the “Plan” phase and going through each subsequent step again.

#### *6.4.2.4 References*

1 https://www.evidentlyai.com/ml-in-production/model-monitoring#model-monitoring-metrics

# Appendix A: Manual for Using the Smart Survey Maturity Model

## A.1 Objective of this user manual

The objective of this user manual is to give guidance to use the smart survey maturity model (as described in this report). Using the maturity model is part of an assessment process. First, this assessment process is described. Secondly is described how to use the model itself and how to 'score' the maturity criteria to determine the maturity level.

Side note: Be aware that this is not a perfect model. In the SSI project we did our best to create a useable model and we applied it in two NSIs to evaluate and to improve the model. The best would be to manage this model by a central authority based on user experiences. However, providing advice on this matter was not in scope of the project.

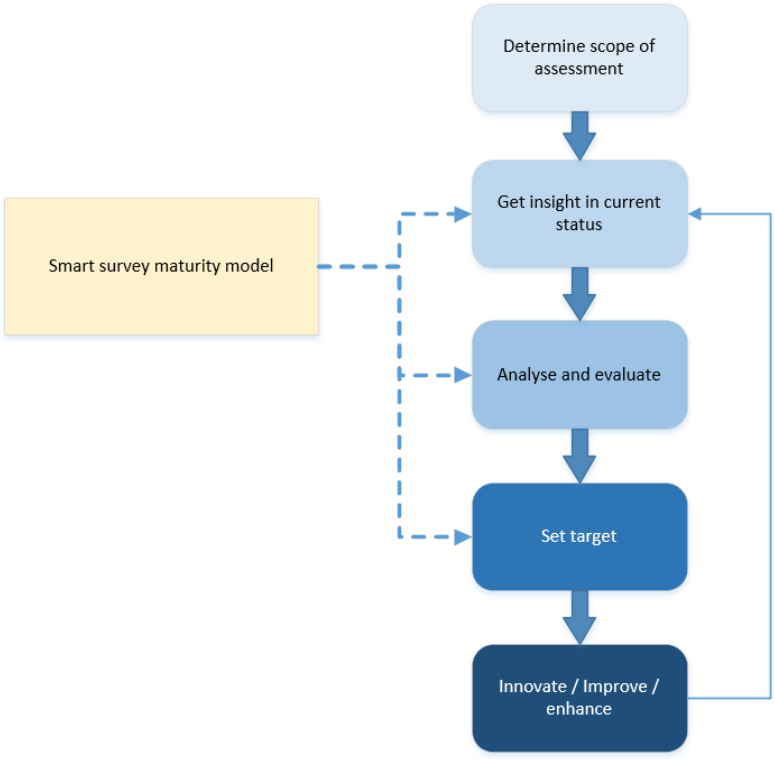
## A.2 Introduction

As described in chapter 3, the maturity model consists of five maturity levels: Awareness, Pilot, Production, Managed and Optimized. And the model consists of five aspects that must be evaluated: Organization, Methodology, IT, Business Process and Legal.

For each aspect and for each maturity level maturity criteria are described (see chapter 3). During the assessment these criteria should be evaluated.

## A.3 Assessment Process

The **assessment process** consists of 5 main steps.




*Assessment process*

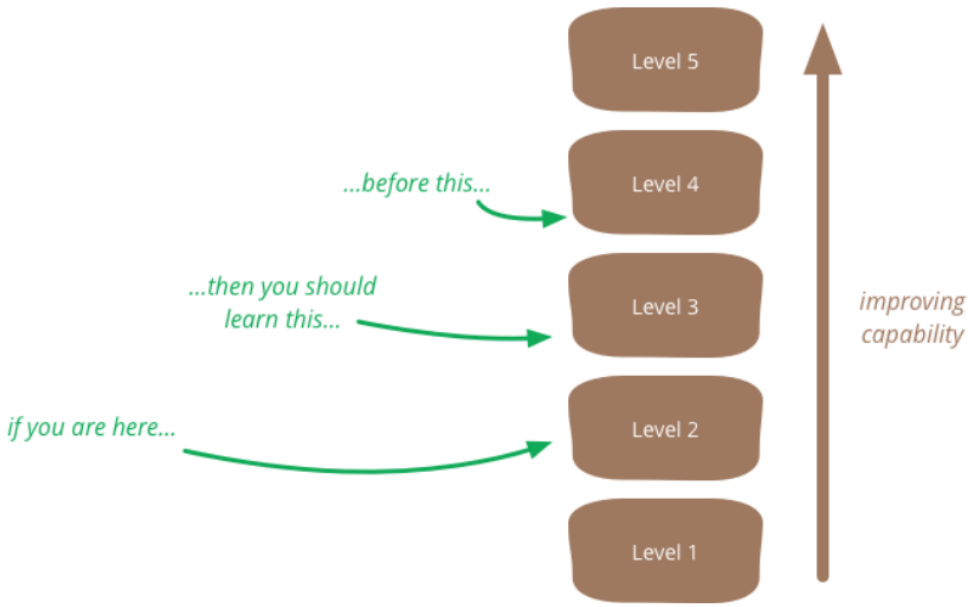
1. First an NSI needs to determine the **scope** of the assessment. The maturity model can namely be used for two scopes: you could determine the maturity of a specific smart solution, or you could determine the maturity of the whole organisation concerning smart solutions.

2. After the scope is determined the second step is to get insight into the **current status** regarding smart solutions. For this the maturity model should be used to evaluate the maturity criteria. Based on this the maturity level for each of the five aspects will be determined.

3. If the current status is determined then the NSI should **analyse** and **evaluate** the results of the findings. So, what is the conclusion that can be drawn? E.g. the conclusion could be that one of the five aspects fell behind in maturity. Another conclusion could be e.g. that a specific part of the process needs extra attention, e.g. the 'design'-phase or the 'data collection'-phase[18].

4. After that **targets** can be set. It is not per se that the target is to get overall to a higher maturity level. If the conclusion is that e.g. one of the aspects is less mature than the other aspects, the target can be e.g. to create more balance, so the target would be to put effort in the 'growth' of one of the aspects.

Looking at the maturity criteria of the model, the NSI can get insight in what steps would be logical to conduct. So, the NSI should look, for a specific aspect, to the maturity criteria of the next higher maturity level.




Source: https://martinfowler.com/bliki/MaturityModel.html

5. When targets are set, **improvements** can be made, focused on the targets. Improvement can be the improvement of an existing smart solution in the NSI to reach a 'next level', but it could also be an innovation activity where a new type of smart solution is introduced for a survey. But also, smaller enhancements in the current situation can be valuable.

[18] In terms of GSBPM

**Iterative process**

It should be clear that the assessment process is an **iterative** process. It is part of continuous improvement. It is not a daily, weekly or monthly process. Think more of it as a yearly process. The process could give input to annual plans for improving your business processes (in a broader perspective). The smart survey assessment is supportive to the broader statistical business improvement process.

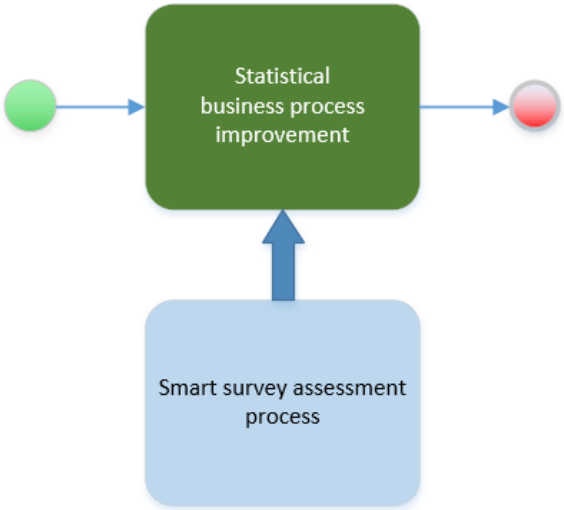



*Illustration: smart survey assessment as part of a larger improvement cycle*

## A.4 How to use the model to assess the current status?

**Preparation**

Best is to make one person responsible for the assessment process. We call this the 'assessment leader'. Each of the five aspects of the maturity model must be assessed by an employee who possesses at least a basic understanding of the specific aspect being evaluated (also referred to as an assessor), ensuring that the evaluation is accurate.

**Baseline Measurement**

The assessment begins at assessing the Baseline. So, check based on the criteria mentioned in the Baseline of the model whether you 'tick' the boxes. If you don't tick all the boxes this should be seen as a warning that probably you are not ready to start using smart solutions and using this model will not add much value. However, this is not hard science. The model is not perfect and if you think that there is no obstacle, of course feel free to use the model and assess your maturity.

**The process**

The process of determining the current maturity status can be done in various ways. Below is a summary of the different methods, along with their main advantages and disadvantages.

- Interviews
  *Required preparation*: Develop interview questions that address the criteria of the smart survey maturity model. Schedule and conduct the interviews with the assessors.
  *Advantage*: You gain detailed information that is often difficult to obtain through surveys.

The interviewee can share insights that help refine the model.
*Disadvantage*: Time-consuming, Results can be subjective and vary depending on the interviewer.

- Surveys
  *Required preparation*: Design a survey with clear, measurable questions that relate to each aspect of the maturity model. Create a customized survey for each assessor, including only the questions they need to answer. Distribute the survey to the assessors and ensure the assessors have enough time to complete it.
  *Advantage*: A quick way to gather data and easy to determine the level based on the collected information.
  *Disadvantage*: Limited in providing contextual information, designing a survey is time-consuming.

- Workshops
  *Required preparation*: Prepare an agenda that covers all the criteria from the aspects of the smart survey maturity model. Facilitate the workshop by encouraging open discussion, allowing everyone to share their insights.
  *Advantage*: Fosters a sense of involvement in the process and helps in reaching consensus on the assessment of the criteria. Probably the least time-consuming method, compared to the above.
- *Disadvantage:* Bringing assessors together at a specific time can be challenging.

- Literature Review
  *Required preparation*: Collecting sources within the NSIs that contain relevant information on various aspects. These can include documents, emails, or websites. Analyse the literature to identify common themes, standards, and best practices. Derive specific criteria that reflect the different levels of maturity. After drafting the criteria, cross-check them with the assessors to ensure accuracy.
  *Advantage:* The assessment of the criteria can be completed largely independently by each assessor
  *Disadvantage:* Time-Consuming, integrating various theories into a coherent set of criteria may require careful analysis.

- Self-assessment
  In this case you also have at least one assessor per aspect. The assessment leader hands over the model to each of those assessors. Each assessor does an assessment purely based on his/her current knowledge of that single aspect. If there are more than one assessors involved in a specific aspect, you need to have a way to get to a single score for the aspect. One way is to ask those 'aspect assessors' to meet and to reach a conclusion, or the assessment leader calculates the average.
  *Required preparation*: Introduce each assessor (high level) into the model and the purpose.
  *Advantage*: Not much time needed.
  *Disadvantage*: There is in principle no communication between the assessors of the different aspects. This can lead to valuable insights being lost.

At Statistics Netherlands an assessment was done with an early draft version of the model. This was done through self-assessment, but afterwards the assessment leader had a short conversation with each assessor to get more context and to evaluate how the assessment process was experienced. Statistics Netherlands experienced this as a good efficient way to do the assessment.

**Scoring**

First, it is good to look at the general characterization of each level. Do you recognize your situation in the characterization of a level? Maybe you will recognize your situation not in one but in two (or even three) levels. Because probably in practice it is the case that some parts of one level are applicable and other parts of another level.

Doing this will give you a general idea of the appropriate level for your situation.

Second, the idea is that you are going to determine the level of maturity per aspect. For each aspect maturity criteria are described per level. The scoring is not hard science; we do not provide a method to get a numerical/quantitative score. Scoring is qualitative. The way to go is to read the criteria per level and then to assess whether those are applicable in your situation. Do the criteria describe the situation you are in? Not all criteria may be applicable in your situation; just leave them out. If your idea is that the criteria do not describe the situation you are in, this could mean two things: you did not yet reach this level of maturity or you already past this level of maturity. To get clear, you have to check the criteria of the next maturity level.

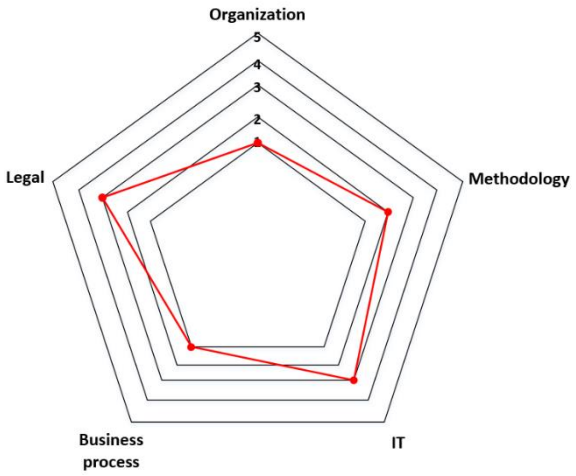
**Outcome**

From the assessment process, the NSI can determine the current maturity level. The assessment results in a level per aspect, reflecting the maturity of each aspect. It could also be that you 'tick' many boxes of a certain level, but also many of an adjacent level. You could then decide that you are 'in between' two levels.

And based on the five different aspect levels, the NSI could (if desired) denominate the overall maturity level[19].

The result could be visualized e.g. using a spider-web.

In this example for each aspect the maturity level is indicated. E.g., Legal is at level 3 and Organization is at level 1. In this case you could say that there is an imbalance in maturity between the different aspects. So, a goal could be e.g. to put effort to get Organization on a more mature level.




[19] We don't provide an algorithm how to do this.

## Appendix B: Process building blocks: Link to GSBPM and BREAL

The goal of this table is to relate each process building block to the phases / sub processes of the GSBPM framework and the business functions of the BREAL[20] framework (see [5]).

| Process building block | GSBPM phase / sub-process | BREAL business function |
|---|---|---|
| **Smart solution design** | | |
| Overarching design of the smart solution | Design: 2.3 Design collection<br><br>Design: 2.6 Design production systems & workflow | Trust management<br><br>Method and tool development for new statistics |
| Decide on involvement of respondent | Design: 2.3 Design collection<br><br>Design: 2.6 Design production systems & workflow | Trust management<br><br>Method and tool development for new statistics |
| Determine reporting period for household | Design: 2.3 Design collection | Method and tool development for new statistics |
| Write DPIA | Design: 2.3 Design collection | Legislative work participation<br><br>Trust management |
| **App** | | |
| Design the app functionality | Design: 2.6 Design production systems & workflow | Method and tool development for new statistics |
| Design the app IT architecture | Design: 2.6 Design production systems & workflow | Method and tool development for new statistics |
| Design app UI/UX | Design: 2.6 Design production systems & workflow | Method and tool development for new statistics |
| Build and test the app | Build: 3.1 Reuse or build collection instruments<br><br>Build: 3.5 Test production systems | Method and tool development for new statistics |
| Deploy the app in the app store | Build: 3.7 Finalize production systems OR<br><br>Collect: 4.2 Set up collection | Deployment |
| Inform respondent how to download and install app | Collect: 4.3 Run collection | Acquisition and recording |
| Download and install app on device | Collect: 4.2 Set up collection | Acquisition and recording |

[20] Big data Reference Architecture and Layers

| | | |
|---|---|---|
| Provide credentials for app | Collect: 4.3 Run collection | Acquisition and recording |
| **Machine-learning model – training & maintenance** | | |
| Choose and configurate machine-learning model structure | Build: 3.1 Reuse or build collection instrument | Modelling and interpretation |
| Collect data to train or update the ML-model | Evaluate: 8.1 Gather evaluation inputs<br>Build: 3.1 Reuse or build collection instrument | Modelling and interpretation |
| Train and test the ML-model | Analyse: 6.2 Validate outputs<br>Build: 3.1 Reuse or build collection instrument | Modelling and interpretation |
| Monitor ML-model performance | Evaluate: 8.1 Gather evaluation inputs<br>Evaluate: 8.2 Conduct evaluation<br>Evaluate: 8.3 Agree on action plan | Quality management<br>Evaluate |
| Update/retrain the ML-model | Build: 3.1 Reuse or build collection instrument | Modelling and interpretation |
| Develop work instructions for manual coding | Build: 3.7 Finalise production system | Quality management |
| **Collection and communication strategy** | | |
| Design the collection and communication strategy | Design: 2.3 Design collection | Manage statistical methodology |
| Ask the respondent for consent | Collect: 4.3 Run collection | Acquisition and recording<br>Trust management |
| Inform respondents about privacy | Collect: 4.3 Run collection | Acquisition and recording<br>Trust management |
| Send a reminder | Collect: 4.3 Run collection | Acquisition and recording |
| **Integrate into a production system** | | |
| Integrate the app channel into the production system | Build: 3.1 Reuse or build collection instruments | Deployment |
| **Providing receipt** | | |

| | | |
|---|---|---|
| Provide a photo of receipt | Collect: 4.3 Run collection | Acquisition and recording |
| Upload e-receipt | Collect: 4.3 Run collection | Acquisition and recording |
| Answer additional questions about the receipt | Collect: 4.3 Run collection | Acquisition and recording |
| Enter receipt information manually | Collect: 4.3 Run collection | Acquisition and recording |
| Store photos and/or uploaded e-receipts | Collect: 4.3 Run collection | Acquisition and recording |
| Scan (bar/QR-) code of the product | Collect: 4.3 Run collection | Acquisition and recording |
| Record the receipt data in the diary | Collect: 4.3 Run collection | Acquisition and recording |
| **OCR/NLP** | | |
| Deploy the microservice 'OCR' | Build: 3.7 Finalise production systems | Deployment |
| OCR/NLP the receipt | Collect: 4.3 Run collection | Data representation |
| Redirect OCR/NLP results to the respondent | Collect: 4.3 Run collection | Data representation |
| Edit OCR/NLP results | Collect: 4.3 Run collection | Data representation |
| **Geolocation** | | |
| Deploy the microservice 'Geolocation' | Build: 3.7 Finalise production systems | Deployment |
| Collect geolocation points | Collect: 4.3 Run collection | Data representation |
| Derive motions and stops | Collect: 4.3 Run collection | Data representation |
| Add context to stops | Collect: 4.3 Run collection | Data representation |
| Edit Geolocation results | Collect: 4.3 Run collection | Data representation |
| **Suggester lists** | | |
| Maintain / update suggester list | Build: 3.1 Reuse or build collection instruments | Method and tool development for new statistics |
| **Diary** | | |
| Link data from 'person level' to 'diary level' | Collect: 4.3 Run collection | Acquisition and recording |
| Determine the moment of response for diary data | Collect: 4.3 Run collection | Acquisition and recording |

| | | |
|---|---|---|
| **Interviewer** | | |
| Train interviewers | Collect: 4.2 Setup collection | Human resource management<br>Quality management |
| Visit respondent by the interviewer | Collect: 4.3 Run collection | Acquisition and recording |
| **Data processing** | | |
| Make data available for data processing | Collect: 4.4. Finalise collection | - |
| Integrate data | Process: 5.1 Integrate data | Integrate survey and register data<br>Data wrangling |
| Design rules for integrating data | Design: 2.5 Design processing and analysis | |
| **Classifying** | | |
| Classify automatically | Process: 5.2 Classify and code | - |
| Classify manually | Process: 5.2 Classify and code | - |
| Make photos / e-receipts available for manual coding | Collect: 4.3 Run collection | - |
| Edit classification by respondent | Collect: 4.3 Run collection | |
| **Helpdesk** | | |
| Train helpdesk employees | Collect: 4.2 Set up collection | Support statistical production<br>Human resource management<br>Quality management |
| Answer questions from respondent | Collect: 4.3 Run collection | Support statistical production |
| **Monitoring and analysing** | | |
| Monitoring process | Evaluate: 8.1 Gather evaluation inputs<br>Evaluate: 8.2 Conduct evaluation<br>Evaluate: 8.3 Agree on action plan | Quality management<br>Evaluate |
| Analyse app usage | Evaluate: 8.1 Gather evaluation inputs<br>Evaluate: 8.2 Conduct evaluation | Quality management<br>Evaluate |

| | | |
|---|---|---|
| | Evaluate: 8.3 Agree on action plan | |
| Analyse usage of smart solution | Evaluate: 8.1 Gather evaluation inputs<br>Evaluate: 8.2 Conduct evaluation<br>Evaluate: 8.3 Agree on action plan | Quality management<br>Evaluate |
| Analyse mode effects | Evaluate: 8.1 Gather evaluation inputs<br>Evaluate: 8.2 Conduct evaluation<br>Evaluate: 8.3 Agree on action plan | Quality management<br>Evaluate |
| **Provide respondent with smart device** | | |
| Create manual how to install / use the device | Build: 3.1 Reuse or build collection instruments | Method and tool development for new statistics |
| Inform respondent about privacy | Collect: 4.3 Run collection | Acquisition and recording<br>Trust management |
| Write or extend DPIA | Design: 2.3 Design collection | Legislative work participation<br>Trust management |
| Ship device to respondent | Collect: 4.3 Run collection | Acquisition and recording |
| Interviewer delivers device | Collect: 4.3 Run collection | Acquisition and recording |
| Get device back | Collect: 4.3 Run collection | Acquisition and recording |