












Smart Survey Implementation

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

Work package 3

Developing Smart Data Microservices

Deliverable 3.3: Demonstration of the Developed Smart Data Microservices

Version 1.0, 2025-04-17

Prepared by:

Joeri Minnen (hbits, Belgium), Pieter Beyens (hbits, Belgium), Ken Peersman (hbits, Belgium), Enak Cortebeeck (hbits, Belgium), Noël Mingels (CBS, Netherlands), Jonas Klingwort (CBS, Netherlands), Chris Lam (CBS, Netherlands), Tim de Jong (CBS, Netherlands), Yvonne Gootzen (CBS, Netherlands), Marco Puts (CBS, Netherlands), Jerome Olsen (Destatis, Germany), Joël Van Hoorde (Destatis, Germany), Adrian Montag (Destatis, Germany), Miriam Engel (Destatis, Germany), Fabrizio De Fausti (ISTAT, Italy), Claudia De Vitiis (ISTAT, Italy), Marco Terribili (ISTAT, Italy), Francesca Inglese (ISTAT, Italy)

Work package Leader:

Joeri Minnen (hbits, Belgium)
e-mail address : Joeri.Minnen@hbits.io
mobile phone : +32 (0)497 189503

Index

1. General introduction

This deliverable serves as an extra source of information that goes along the Smart advanced report on Developing Smart Data Microservices. The Smart advanced reports the work completed in Work Package 3 (WP3) of the SSI project, which focuses on developing microservices to support Trusted Smart Surveys and demonstrate an end-to-end data collection process.

Three microservices were developed:

- Receipt Scanning Microservice: Includes an OCR module (non-domain-specific) and a COICOP classification module (specific to Household Budget Surveys).
- GeoService Microservice: Comprises a Geolocation module (non-domain-specific) and an HETUS classification module (specific to Time Use Surveys).
- Energy Microservice: Restricted to a feasibility setup of study and roll-out

Each microservice has a non-domain-specific component for broad usage and a domain-specific component tailored to specific statistical needs. The non-domain-specific development are integrated to the end-to-end data collection process in connection to a data collection platform. These platforms are the MOTUS data collection platform from hbits (Belgium) and the @HBS platform (to be continued after the end date of the SSI project) from CBS (The Netherlands).

The purpose of this document is, first, to provide information on where to find the code of the developments that were carried out within WP3, and, second, to demonstrate the work by showing the integration of the non-domain parts to the data collection platform. Domain-specific parts can be added later following the same principles.

The development work has been carried out as consortium but with different teams being responsible for the underlying parts. Below a table is presented of these tasks together with the institution to be contacted. hbits was the overall lead of WP3.

Table 1: Developments within the SSI project and the responsible members

Developments			
General Microservice architecture			
General service		Views and components	hbits
Receipt Scanning Microservice			
Non-domain specific		Overall	hbits
	OCR Microservice	Image preprocessing	CBS
		OCR text extraction	CBS
		Document Understanding	hbits
		Integration models	hbits
		Pipeline + API	hbits
		Integration to MOTUS	hbits
		Integration to @HBS	CBS
Domain specific		Overall	Destatis
	COICOP Microservice	CPI-data string matching	Destatis

		ML-model	CBS
		Pipeline	Destatis & CBS
GeoService Microservice			
Non-domain specific		Overall	hbits
	Geolocation Microservice part 1	Stop-track clustering	hbits
		API	hbits
		Integration to MOTUS	hbits
	Geolocation Microservice part 2	Mode of transport	CBS
Domain specific		Overall	ISTAT
	HETUS Microservice	Matching POI description to HETUS	ISTAT
		Prediction ranked HETUS list based on POI, time, duration, back-ground info	ISTAT
		Pipeline	ISTAT
Energy Microservice			
Feasibility setup		Overall	CBS

Contact information:

hbits	info@hbits.io
CBS	r.paulussen@cbs.nl
Destatis	lwr@destatis.de
ISTAT	devitiis@istat.it

2. ESSnet-SSI Github repository

The Github ESSnet-SSI repository is the storage space where WP3 partners keep the project files and track changes in relation to the development work that has been done during the project.

Url: https://github.com/essnet-ssi

This repository contains:

- Source Code: The main files of the project, written in a specific programming language.
- README.md: An introductory file explaining the project's purpose, features, installation instructions, and usage guidelines.
- License: EUPL-v1.2 license (https://eupl.eu/1.2.+en) involving the terms under which the code can be used, modified, and distributed.
- Contributing Guide: Instructions for other developers on how to contribute to the project, including coding standards and pull request procedures.
- Issues: A section for reporting bugs, requesting features, or discussing enhancements.
- Pull Requests: Proposed code changes awaiting review before being merged into the main codebase.
- Branches: Different versions of the project, allowing multiple developers to work on features or fixes simultaneously.
- Commits: A history of changes made to the code, with messages explaining each update.
- Actions: Automated workflows for tasks like testing, building, or deploying the project.
- Wikis and Discussions: Collaborative areas for documentation and community interaction.

As pointed out the software is available free and open source under the EUPL-v.1.2 license. This license has a compatibility clause. It allows the software to be combined with other code licensed under compatible licenses. If you distribute the combined work, you can choose to use the EUPL or one of the compatible licenses.

The EUPL provides the software "as-is" without any warranties. The licensor is not liable for any damages resulting from the use of the software.

By this the ESSnet-SSI repository provides a structured environment for version control, collaboration, and open-source contribution. Although no active support can be anticipated it should be in the interest of NSIs to continue working with the developed software, and to share new developments in the life span of this repository.

More information is available in the README.md file in the archive.

3. Demonstrations

During the SSI project, multiple demonstrations were conducted to showcase the development and integration of various microservices. These demonstrations were recorded and are accessible via OpenSocial. They provided insights into the progress, functionality, and implementation of the microservices within the end-to-end data collection process.

First Informational Meeting (October 20th, 2023)

This session presented the status of three models for the OCR microservice and discussed strategies to combine them effectively. The integration of the OCR microservice with relevant platforms was also explored.

- Demonstration by hbits and CBS: https://cros.ec.europa.eu/book-page/information-session-october-2023
- Presentation Slides: https://cros.ec.europa.eu/book-page/information-session-october-2023

Second Informational Meeting (March 22nd, 2024)

The operational OCR microservice and its integration into an independent data collection platform were demonstrated, alongside the COICOP microservice. The session covered the components, platform integration, and AI/ML training strategies.

- Demonstration by hbits, CBS and Destatis: https://cros.ec.europa.eu/book-page/information-session-march-2024-wp3
- Presentation Slides: https://cros.ec.europa.eu/book-page/information-session-march-2024-wp3

Third Informational Meeting (July 5th, 2024)

This presentation focused on the GeoService microservice, including non-domain components like the stop-track and mode of transport models. ISTAT presented work on HETUS classification predictions, and the UI/UX integration in MOTUS was showcased.

- Demonstration by hbits: https://cros.ec.europa.eu/book-page/information-session-july-2024-wp3
- Presentation Slides: https://cros.ec.europa.eu/book-page/information-session-july-2024-wp3

Fourth Informational Meeting (November 22nd, 2024)

The integration of the Receipt Scanning Microservice with a large field test in Germany was presented.

- Demonstration by University of Mannheim, Destatis and hbits: https://cros.ec.europa.eu/book-page/information-session-november-2024-wp2
- Presentation Slides: https://cros.ec.europa.eu/book-page/information-session-november-2024-wp2

Fifth Informational Meeting (February 21st, 2025)

This presentation has the focus on the @CBS application.

- Demonstration by CBS: https://cros.ec.europa.eu/dashboard/trusted-smart-surveys and look for session February 2025

- Presentation Slides: https://cros.ec.europa.eu/dashboard/trusted-smart-surveys and look for session February 2025

NTTS conference (March 12nd, 2025)

This presentation provides an overview of the work performed during the SSI-project.

- Demonstration by hbits as overall WP3 leader: https://cros.ec.europa.eu/ntts2025 and https://webcast.ec.europa.eu/ntts-2025-conference-gasp-2025-03-12
- Presentation Slides: https://cros.ec.europa.eu/ntts2025

SSI end conference (April 3-4th, 2025)

Overall presentation WP3 of technical solutions, and specific for HBS and TUS.

- Demonstration by hbits, CBS, Destatis and ISTAT: https://cros.ec.europa.eu/dashboard/trusted-smart-surveys and look for SSI conference Heerlen April 2025
- Presentation Slides: https://cros.ec.europa.eu/dashboard/trusted-smart-surveys and look for SSI conference Heerlen April 2025