Smart Survey Implementation

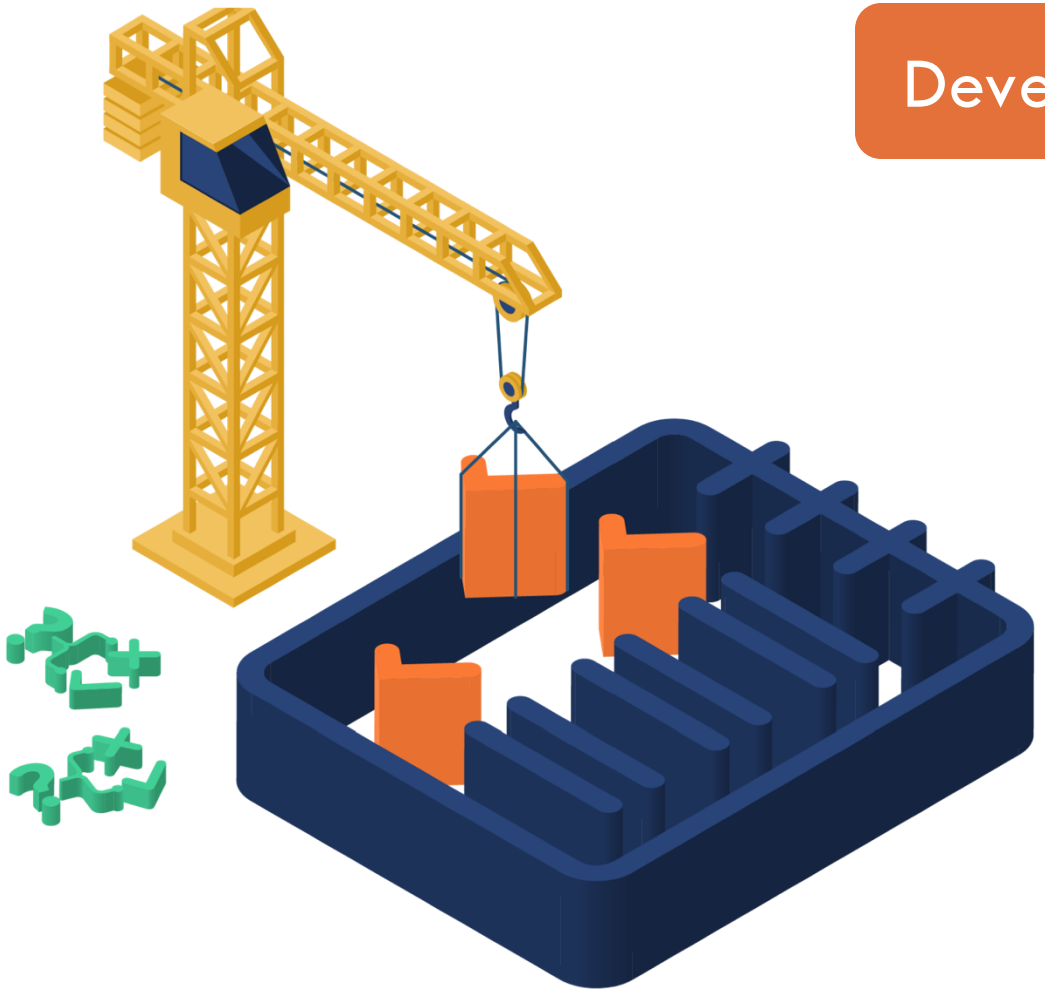
Grant Agreement Number: 101119594 (2023-NL-SSI)

Development and Integration of Microservices

Wednesday, 2025-03-11

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

Goals WP3



Development and Integration of Microservices

For Smart End-to-End Surveys

HBS TUS

Joeri Minnen - hbits

Jerome Olsen - Destatis

Claudia De Vitiis - ISTAT

European funded project to leverage Smart Survey Implementation (SSI): 8 countries

1



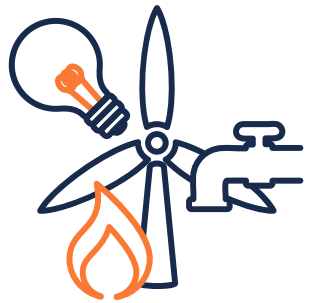
Receipt Scanning

2



GeoService

3



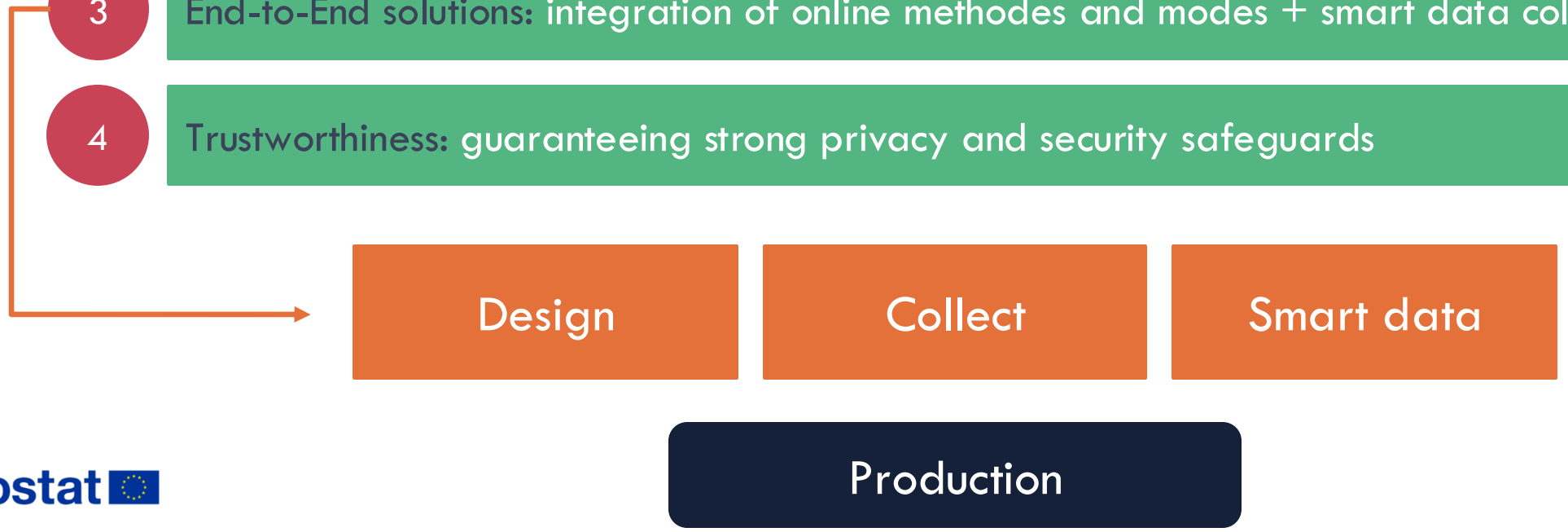
Energy usage

3 developments, one reusable and shareable architecture

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

Road to production: challenges, stakeholders and innovation

1. Citizens: role as active contributors and stakeholders
2. Institutions: role as data collectors, funders and stakeholders
3. End-to-End solutions: integration of online methodes and modes + smart data collection solutions
4. Trustworthiness: guaranteeing strong privacy and security safeguards



eurostat

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

Non-Domain -and- Domain specific components for diary studies

First solution for HBS

Non-domain

Receipt Scanning



X

Domain

COICOP classification



Second solution for TUS

Non-domain

GeoService



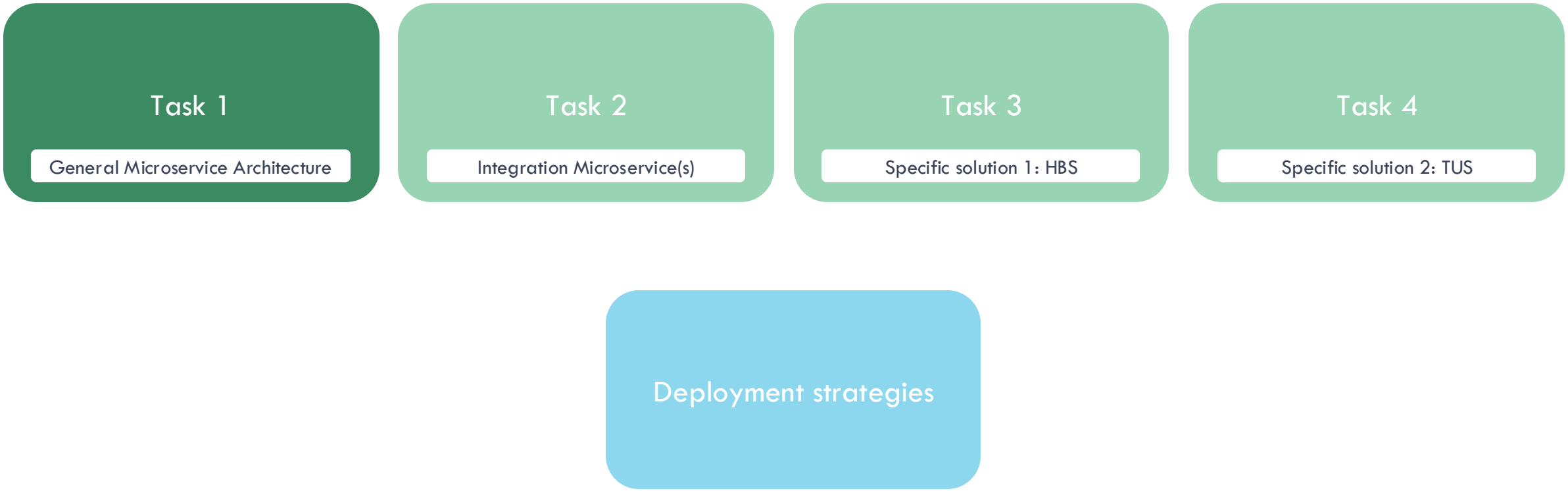
X

Domain

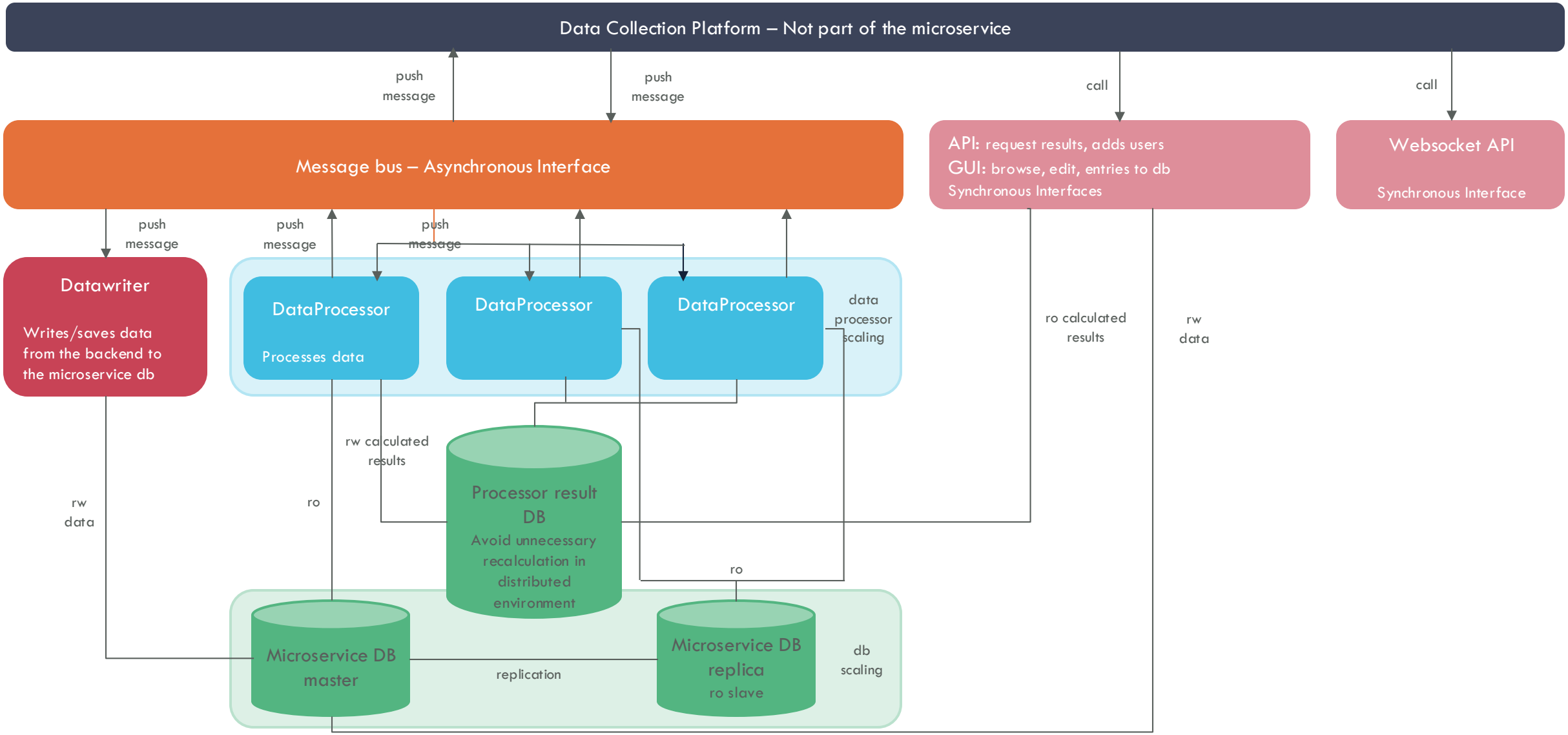
HETUS classification



Task 1: General Microservice Architecture



Microservice architecture and Integration to (any) data collection platform(s)



Task 2: Integrate microservice to achieve End-to-End Solution

Task 1

General Microservice Architecture

Task 2

Integration Microservice(s)

Task 3

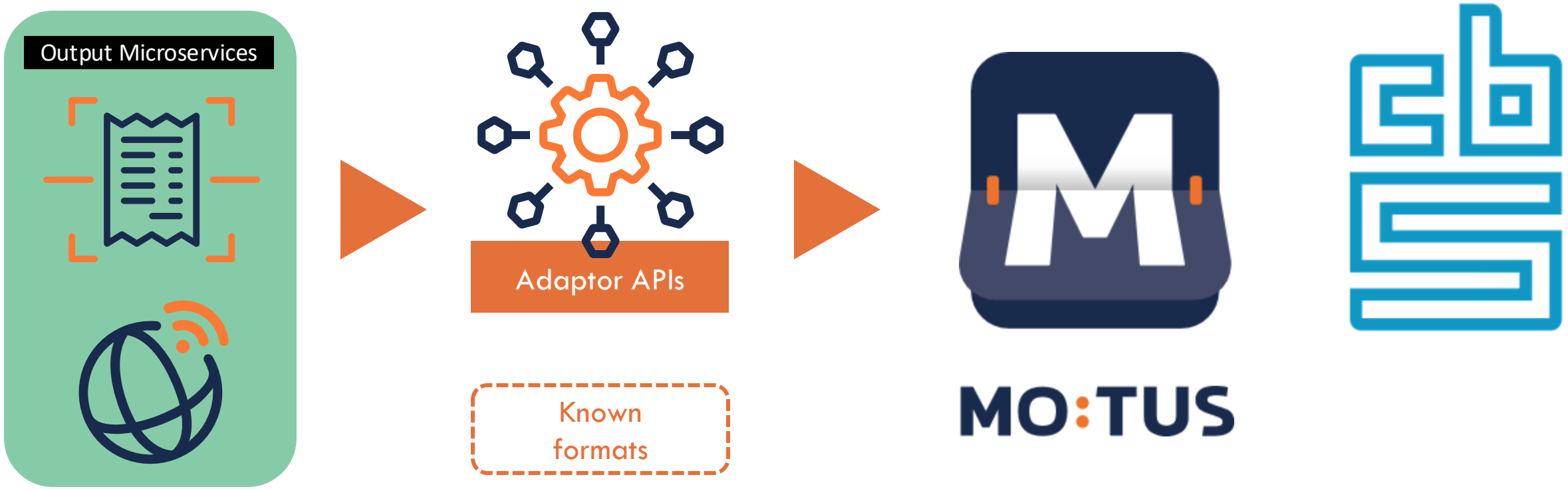
Specific solution 1: HBS

Task 4

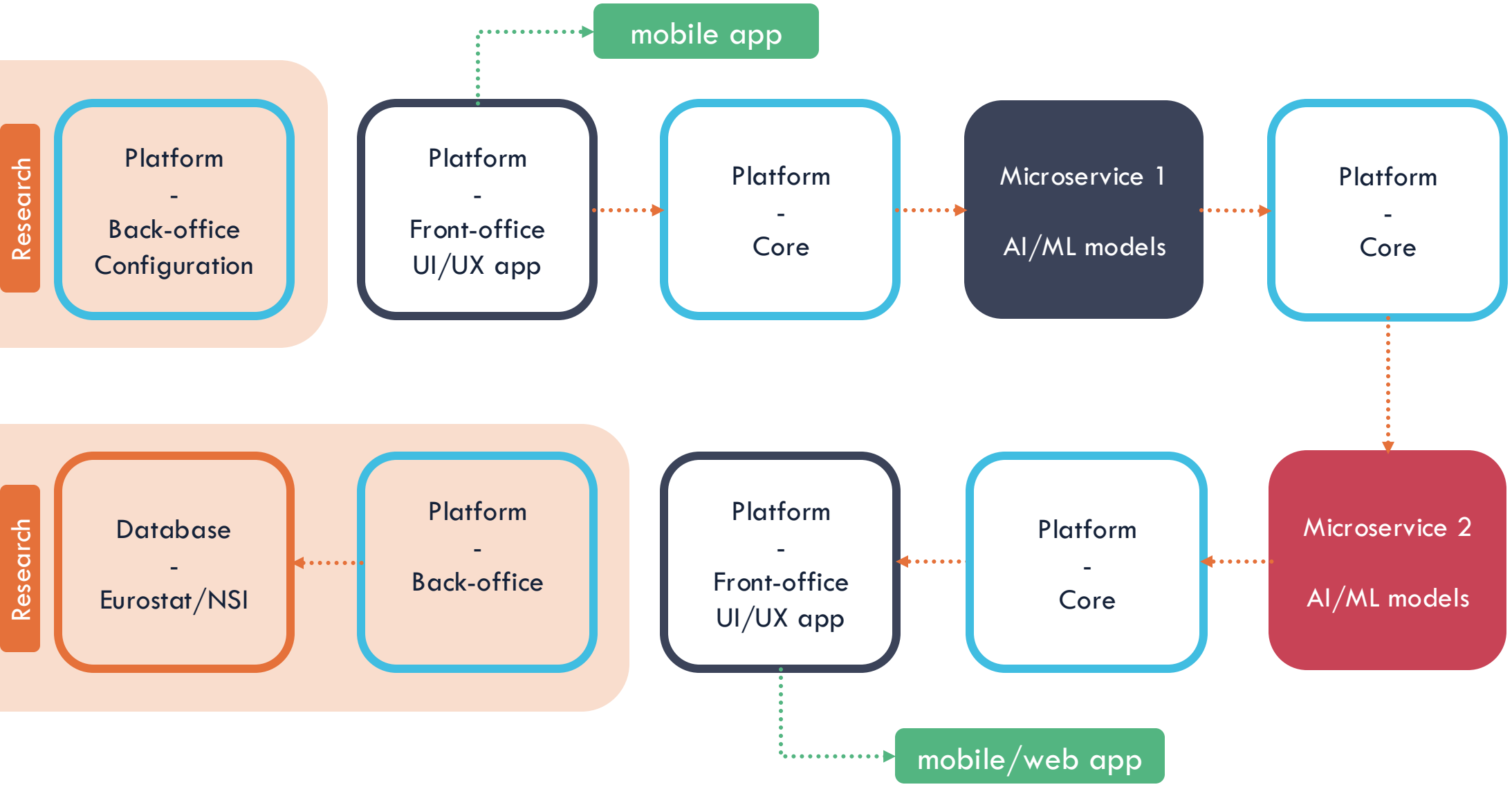
Specific solution 2: TUS

Deployment strategies

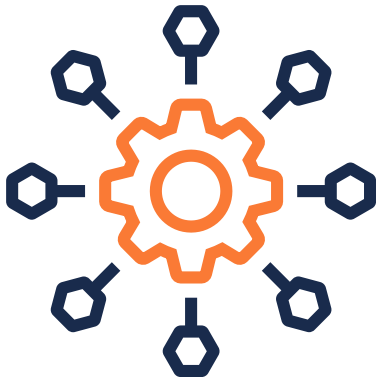
Integration in 2 SSI platforms



End-to-End solution: integrating microservices to become smart



Smart Data = Tentative data



Any type of smart data

User side

Tentative data
(measured data)

Committed data

Integration

UI/UX

Task 3: Specific solution for HBS

Task 1

General Microservice Architecture

Task 2

Integration Microservice(s)

Task 3

Specific solution 1: HBS

Task 4

Specific solution 2: TUS

Deployment strategies

Recording: scan + back-office (non-domain)



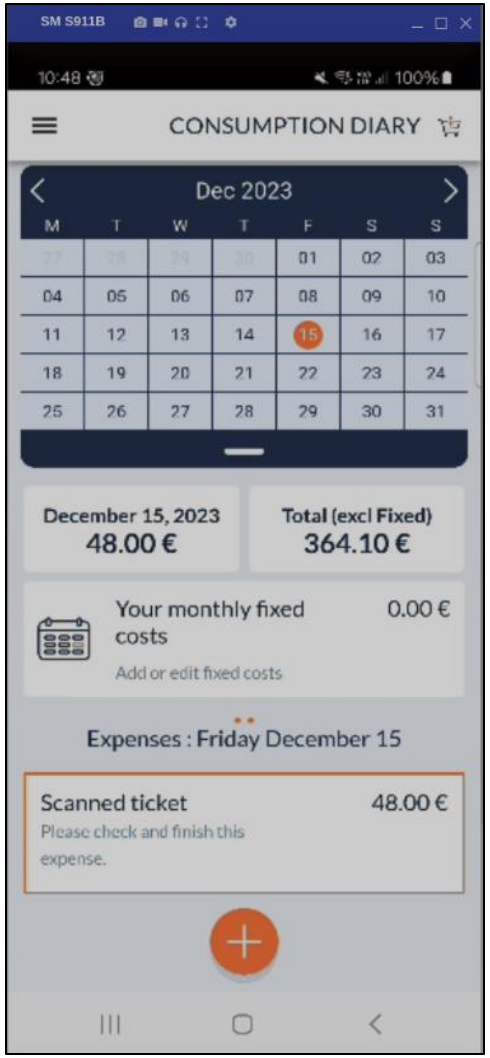


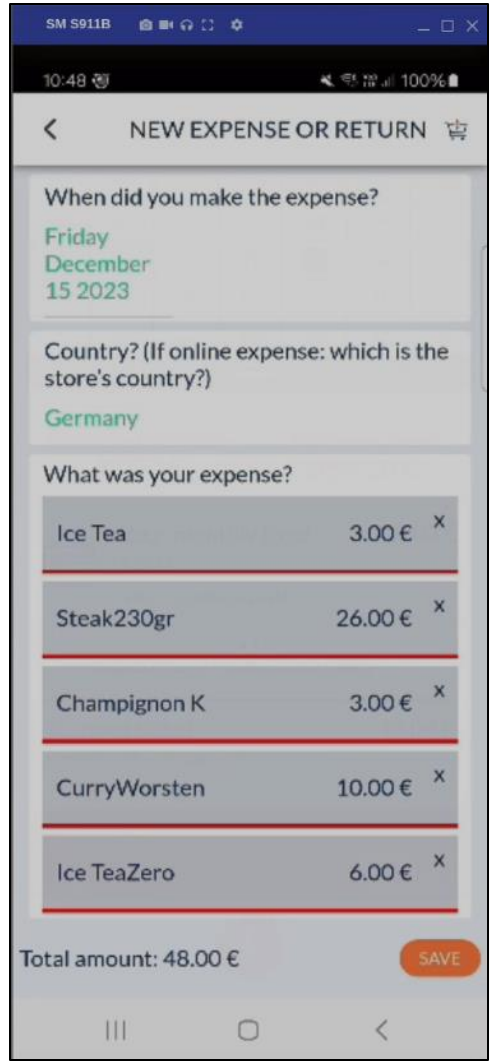
Boxing text

1. OCR

2. Document understanding

Understanding text





COICOP pipeline (domain)

Product descriptions from receipt

COICOP code

ATT1: Data corpus with data linked to COICOP-code

ATT2: Data corpus can be limited to specific domain(s)

Microservice

Pipeline

Receipt data

Matching/ training data







1. String matching techniques

2. Matching learning algorithms



Product ID matching

Str_which (exact)

Stringsim (similarity score)

Str_split (tokenized)

N-gram w/ logic regression

N-gram w/ radom forest

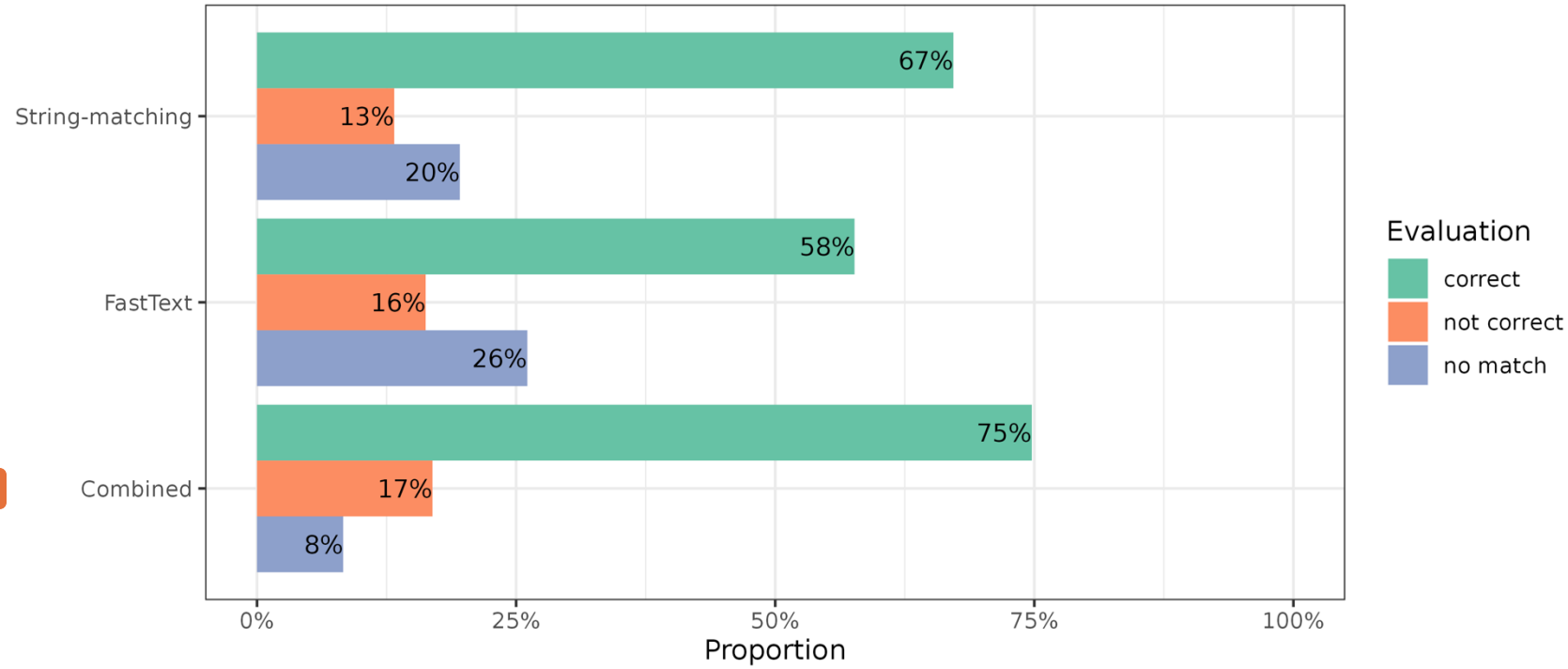
Fast text

Post-survey

3. Manual coding by user

4. Post-coding by coders

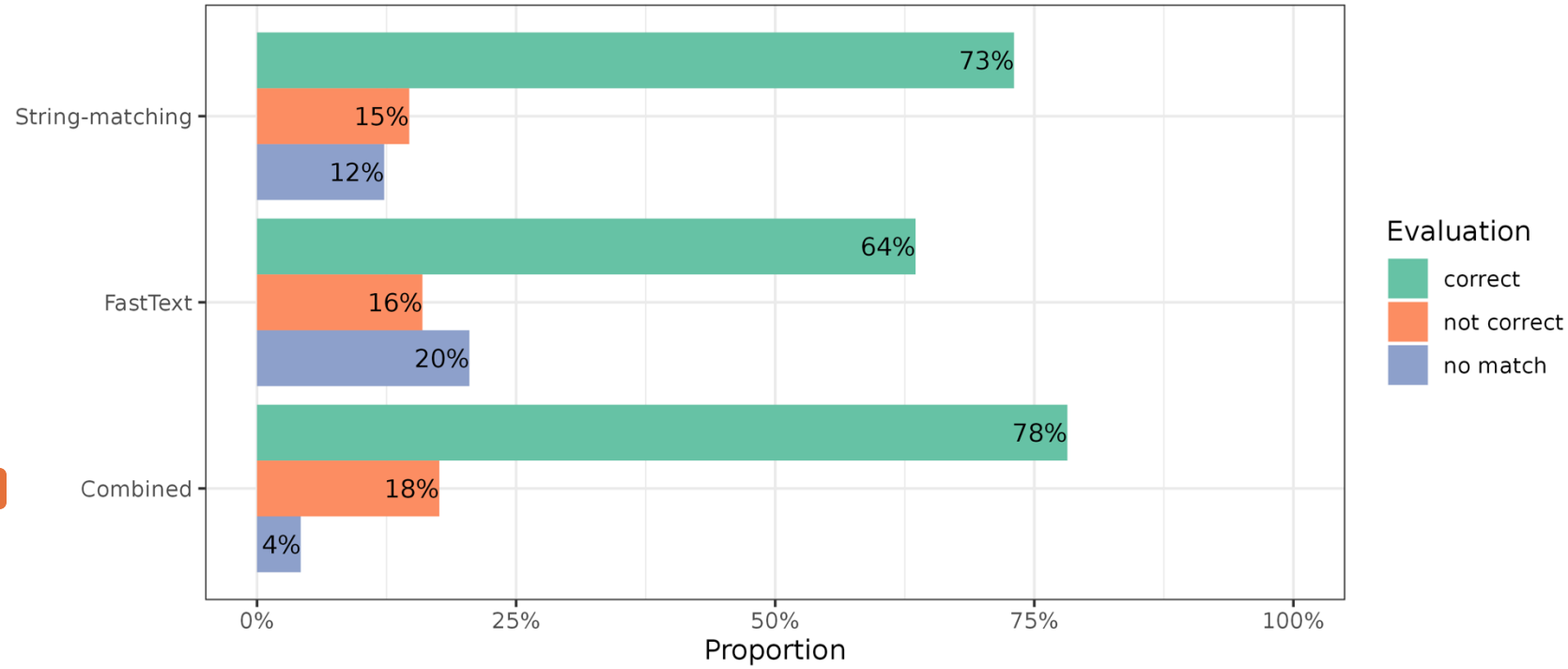
COICOP pipeline results with uncorrected strings (OCR errors)



Pipeline

Text with N = 2,274 product descriptions extracted from receipts

COICOP pipeline results with corrected strings (no OCR errors)

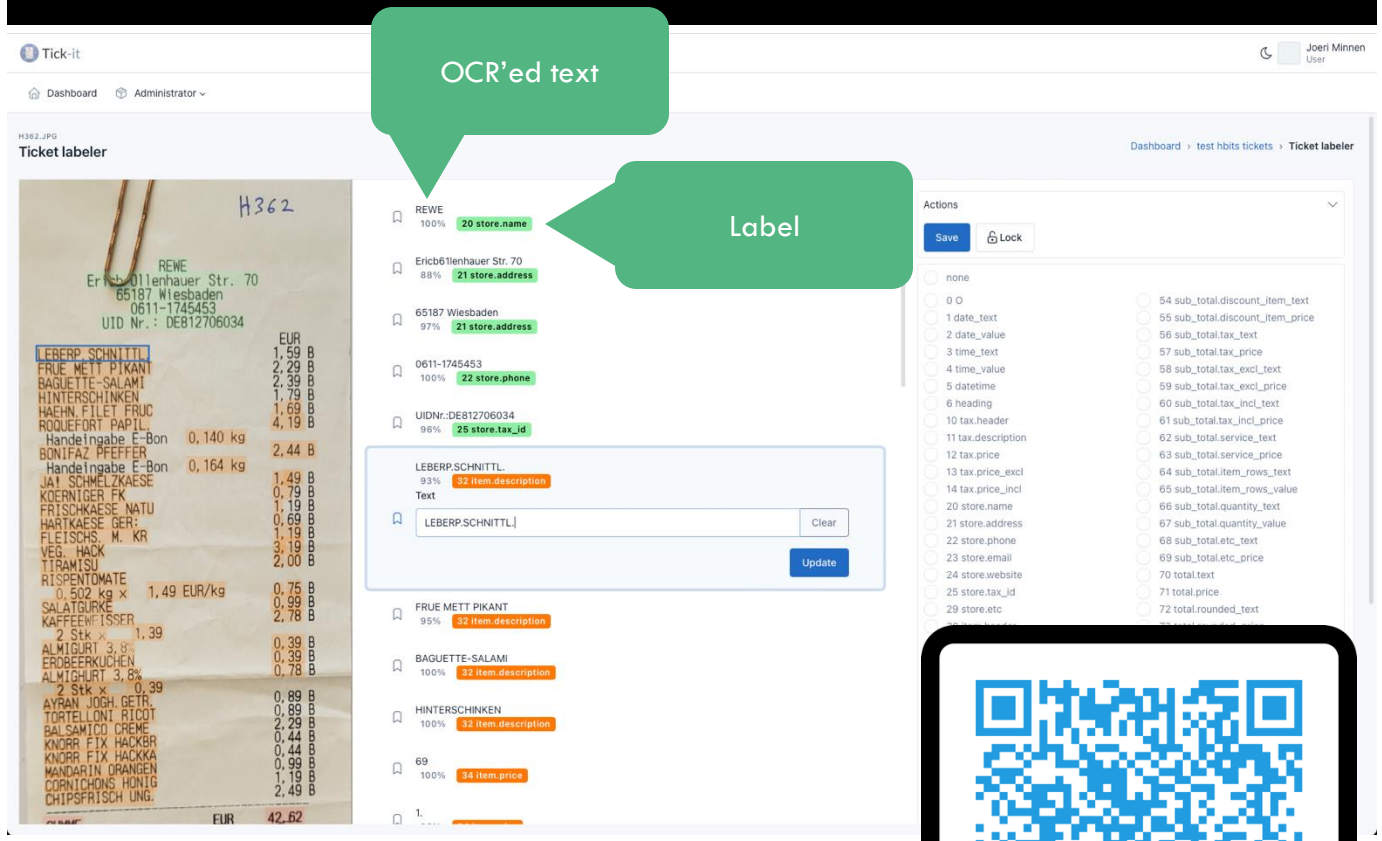


Pipeline

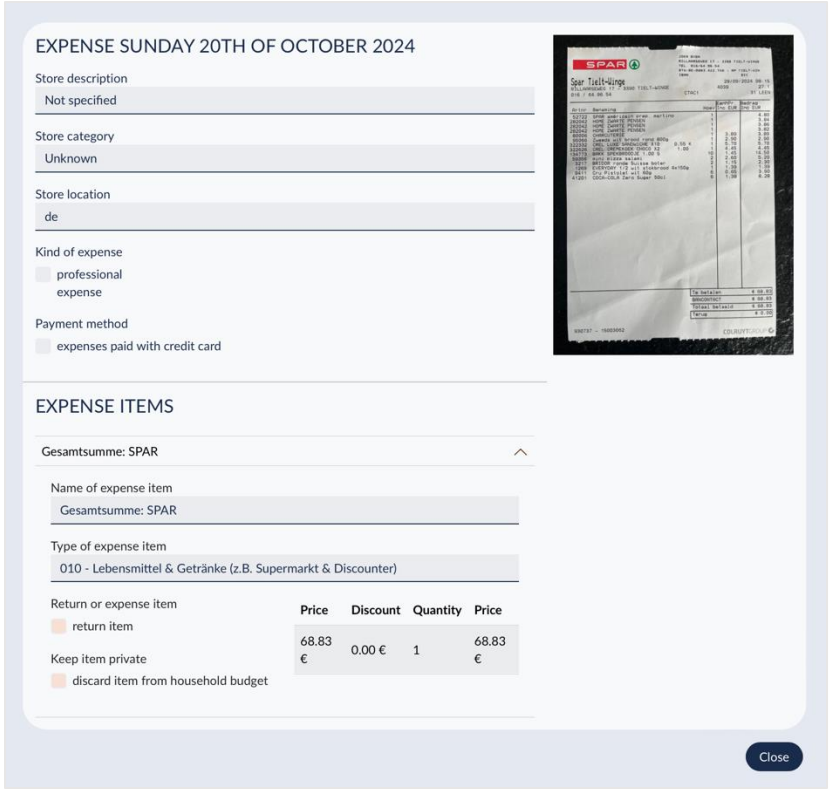
Text with N = 2,274 product descriptions extracted from receipts

Recording: scan + back-office (training non-domain)

Ticket-it – extending the pipeline to a training environment to create trained data or update data to new realities



Researchers view



Need for training data

Task 4: Specific solution for TUS

Task 1

General Microservice Architecture

Task 2

Integration Microservice(s)

Task 3

Specific solution 1: HBS

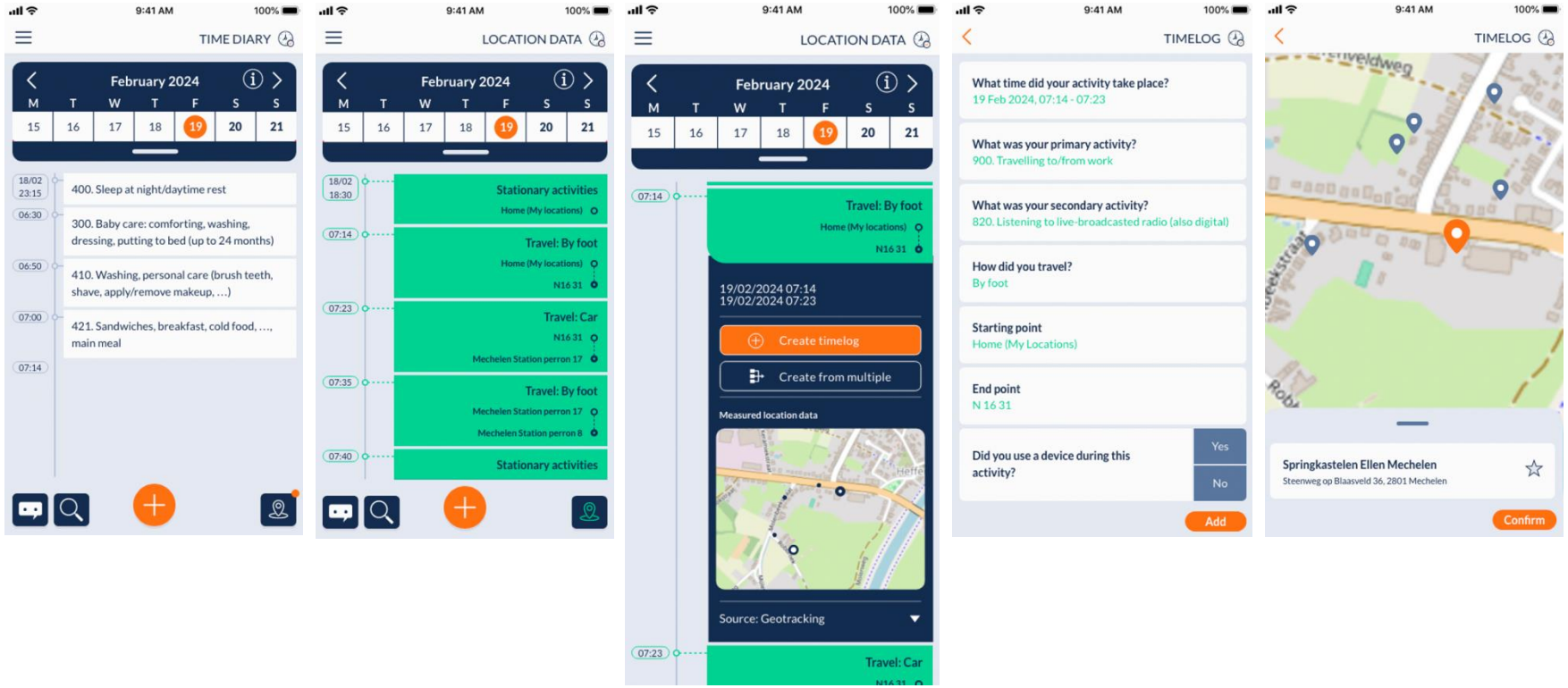
Task 4

Specific solution 2: TUS

Deployment strategies

Recording: geopoints + back-office (visualisation domain/non-domain)

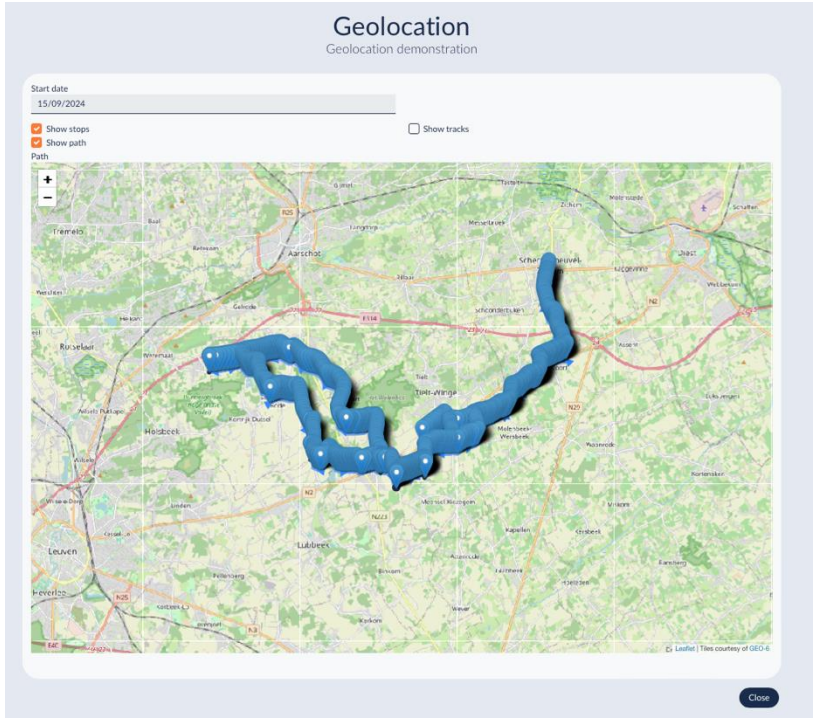
Timeline | Measured data | Detailed view | Adding context | Map visualisation



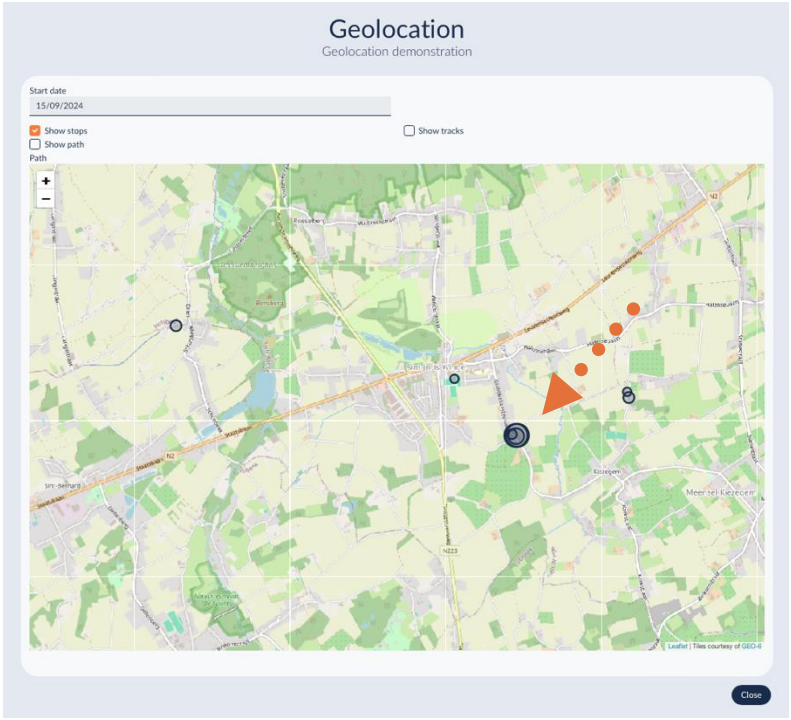
Geopoints + back-office (non-domain)

Stop/track: ATS-OPTICS model

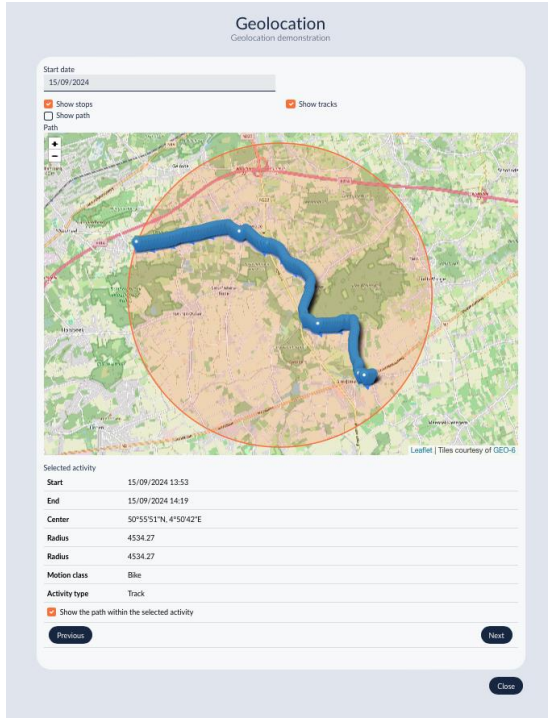
Geopoints per day



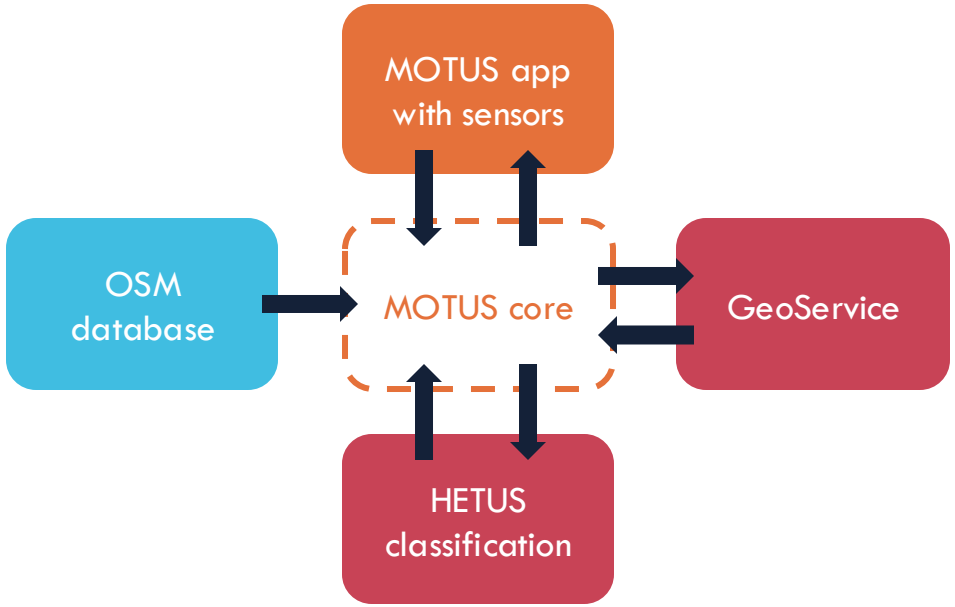
Stop clusters



Track cluster



HETUS pipeline (domain)





1

GPS points (blue)

POIs (green)



2

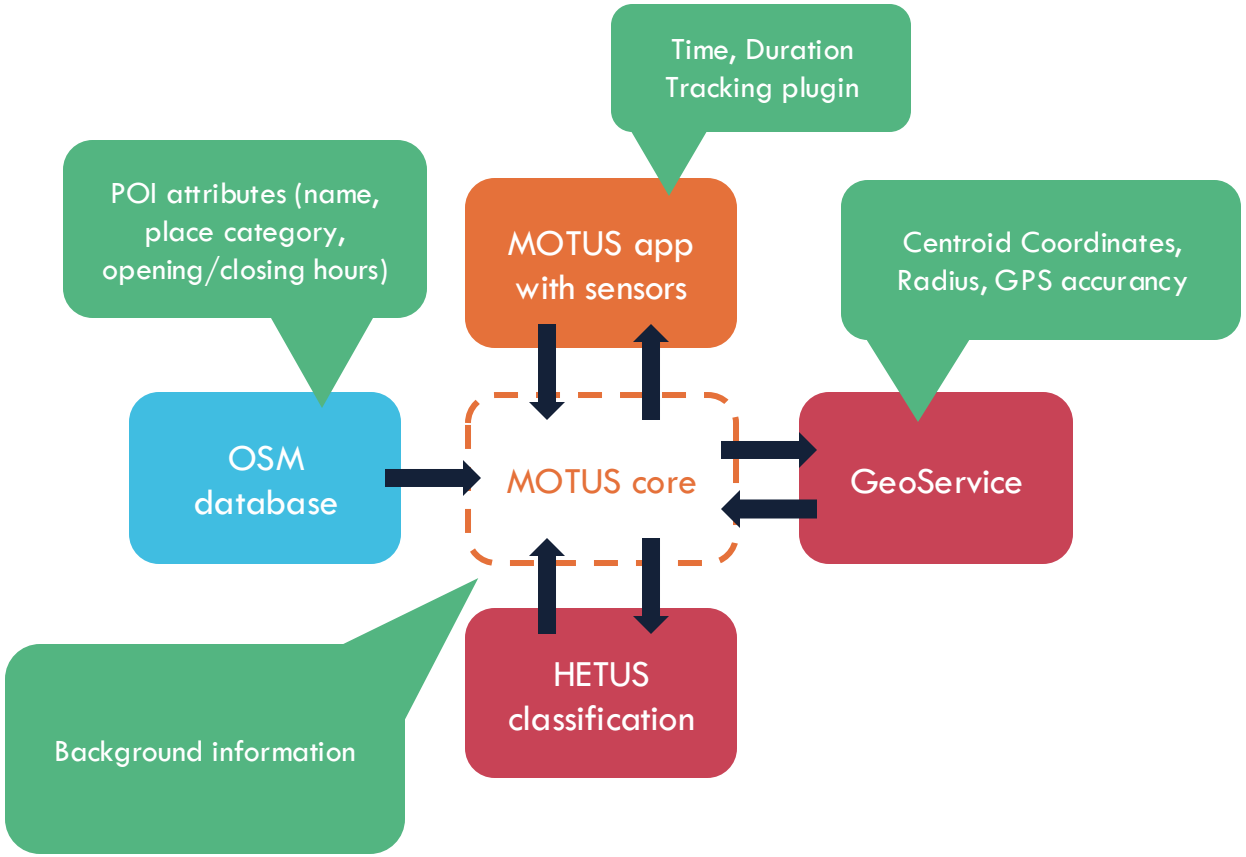
POIs short list (red)

HETUS	ActivityScore	Descr
021	7.400402e-02	021 Eating
361	4.739460e-02	361 Shopping (including online/ e -sho
519	3.842740e-02	519 Other or unspecified social life
032	3.437884e-03	032 Personal care servi ces
732	1.588585e-03	732 Parlour games and play
513	1.237589e-03	513 Celebrations
821	1.227424e-03	821 Watching TV, video or DVD
522	3.925254e-04	522 Theatre and concerts
343	3.686005e-04	343 Caring for pets
831	3.219861e-04	831 Listening to radio or recordings
383	1.547464e-04	383 Reading, playing and talking with c
811	8.229526e-05	811 Reading periodicals

3

Stop
x
ordered list HETUS

HETUS pipeline (domain)





1

GPS points (blue)

POIs (green)



2

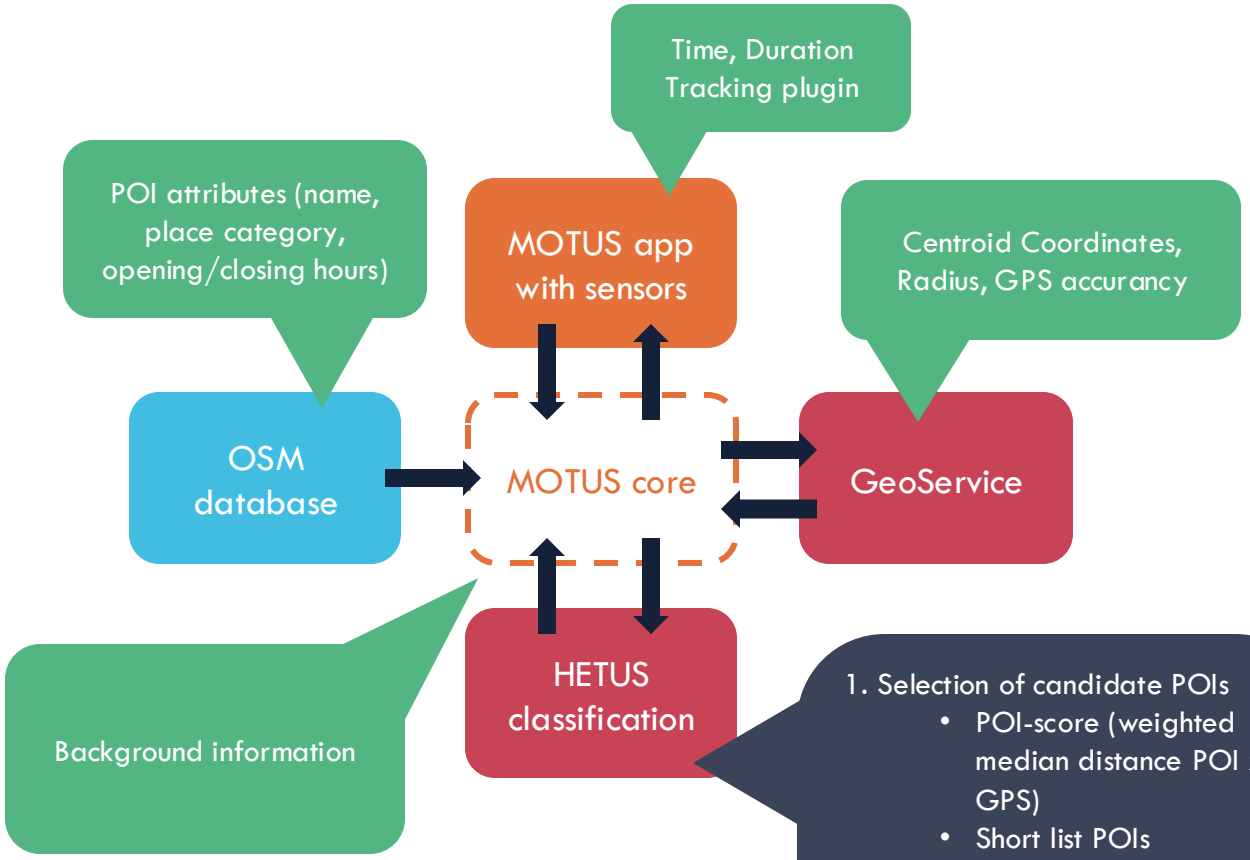
POIs short list (red)

HETUS	ActivityScore	Descr
021	7.400402e-02	021 Eating
361	4.739460e-02	361 Shopping (including online/ e -sho
519	3.842740e-02	519 Other or unspecified social life
032	3.437884e-03	032 Personal care servi ces
732	1.588585e-03	732 Parlour games and play
513	1.237589e-03	513 Celebrations
821	1.227424e-03	821 Watching TV, video or DVD
522	3.925254e-04	522 Theatre and concerts
343	3.686005e-04	343 Caring for pets
831	3.219861e-04	831 Listening to radio or recordings
383	1.547464e-04	383 Reading, playing and talking with c
811	8.229526e-05	811 Reading periodicals

3

Stop
x
ordered list HETUS

HETUS pipeline (domain)





1

GPS points (blue)

POIs (green)



2

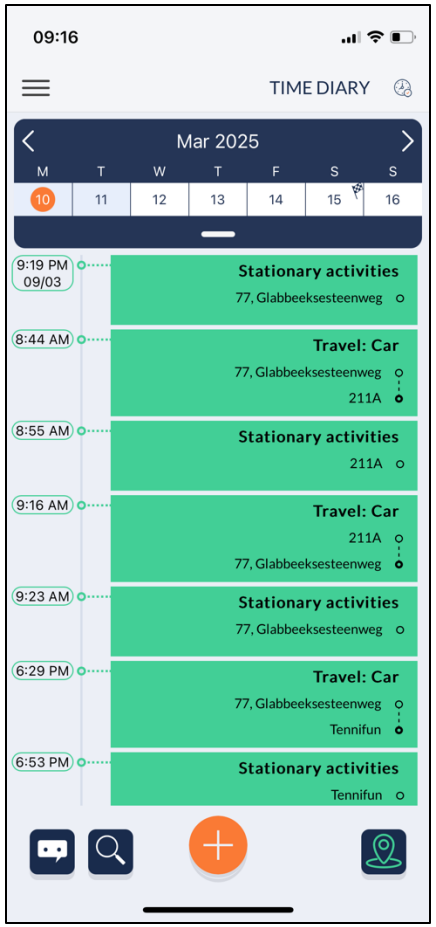
POIs short list (red)

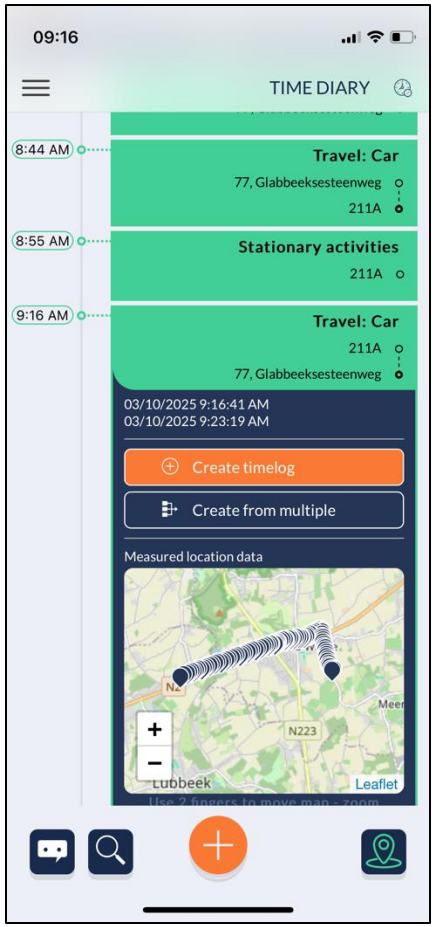
HETUS	ActivityScore	Descr
021	7.400402e-02	021 Eating
361	4.739460e-02	361 Shopping (including online/ e -sho
519	3.842740e-02	519 Other or unspecified social life
032	3.437884e-03	032 Personal care servi ces
732	1.588585e-03	732 Parlour games and play
513	1.237589e-03	513 Celebrations
821	1.227424e-03	821 Watching TV, video or DVD
522	3.925254e-04	522 Theatre and concerts
343	3.686005e-04	343 Caring for pets
831	3.219861e-04	831 Listening to radio or recordings
383	1.547464e-04	383 Reading, playing and talking with c
811	8.229526e-05	811 Reading periodicals

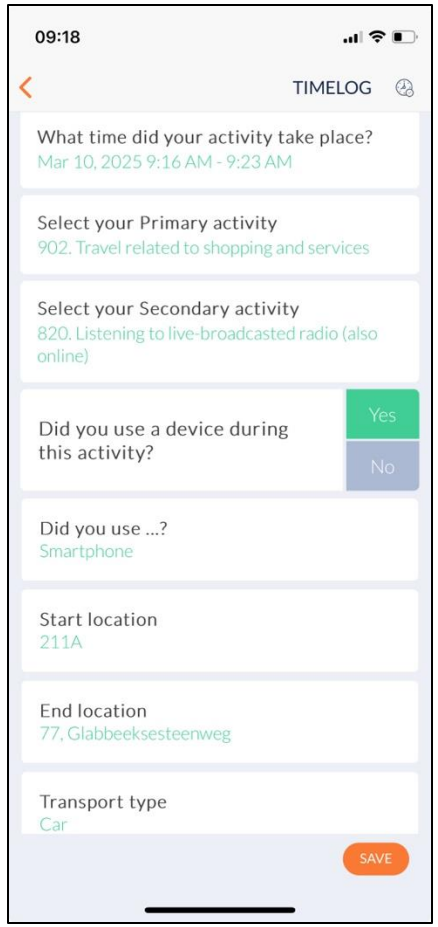
3

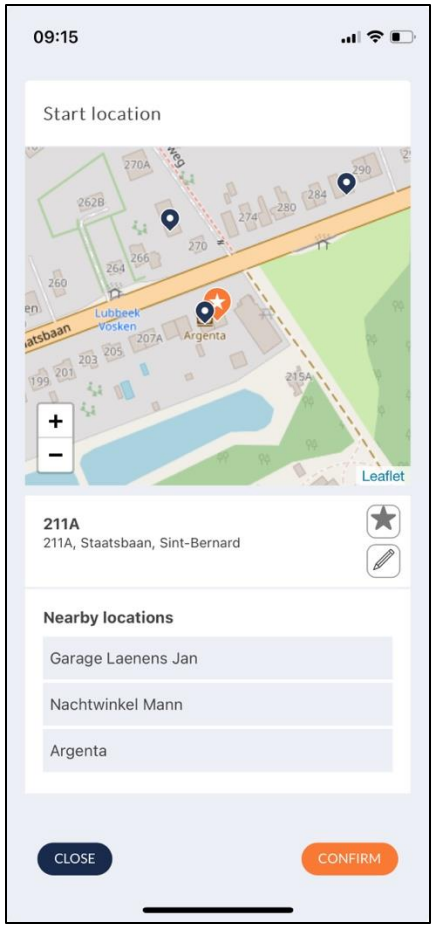
Stop
x
ordered list HETUS

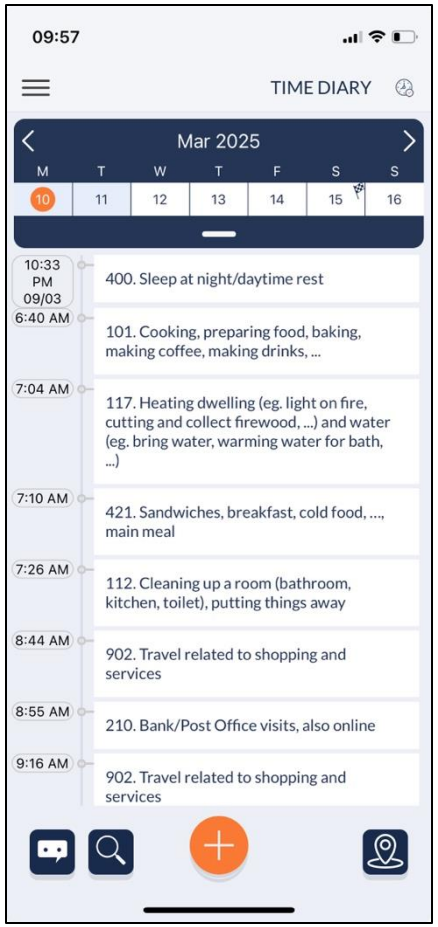
Front-office - March 10th – Car drive from home to bank (service)





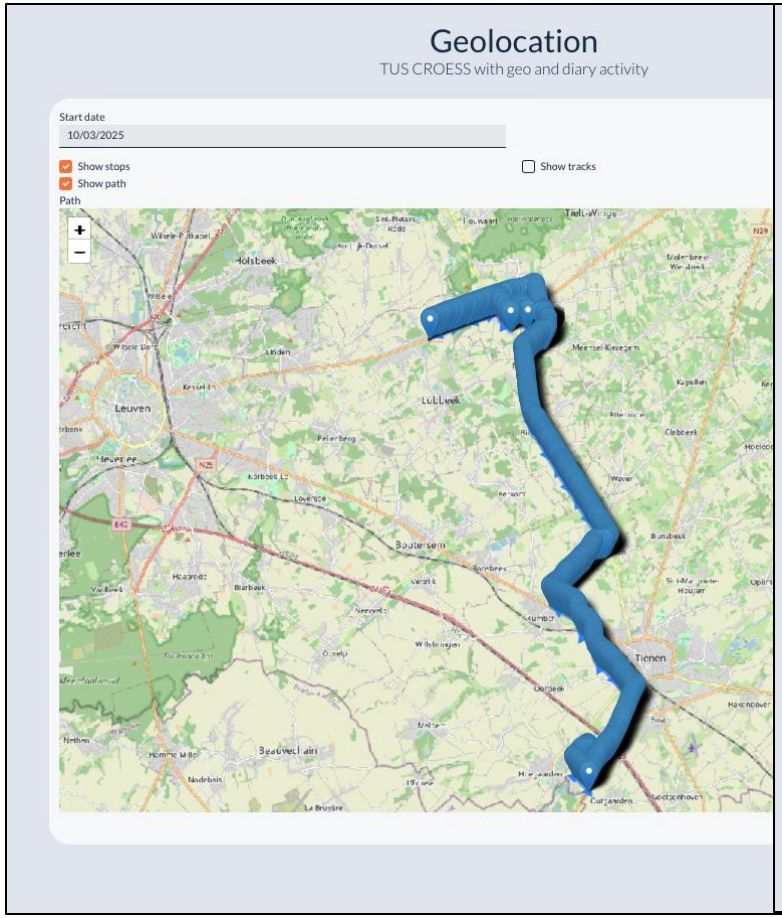






Back-office - March 10th – Car drive from home to bank (service)

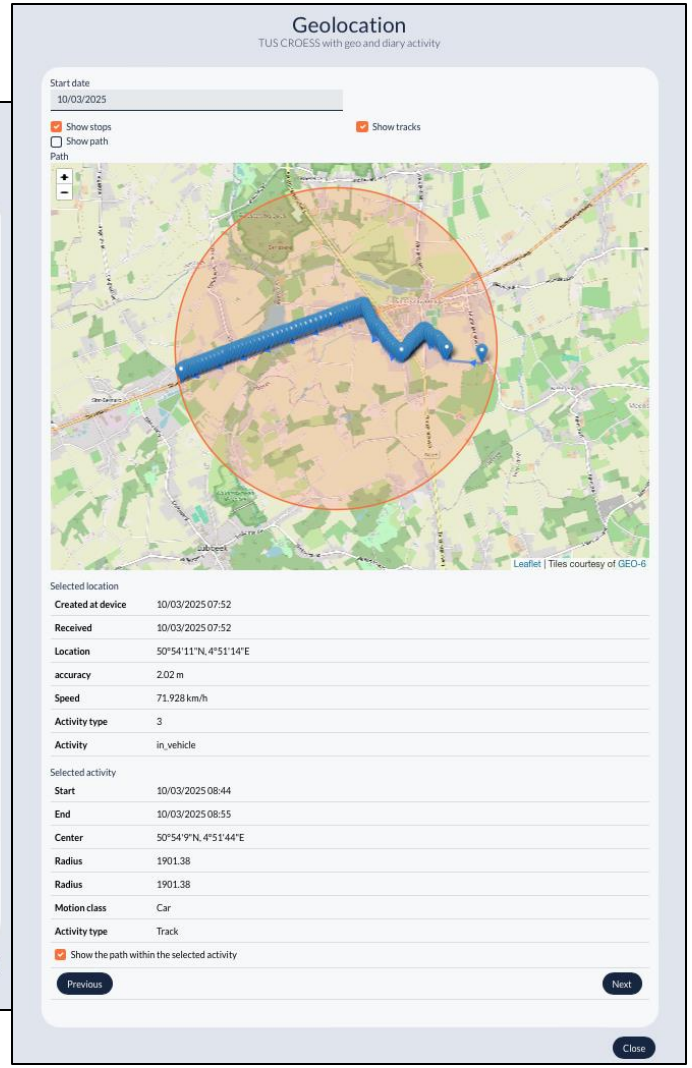
Geopoints 10.03



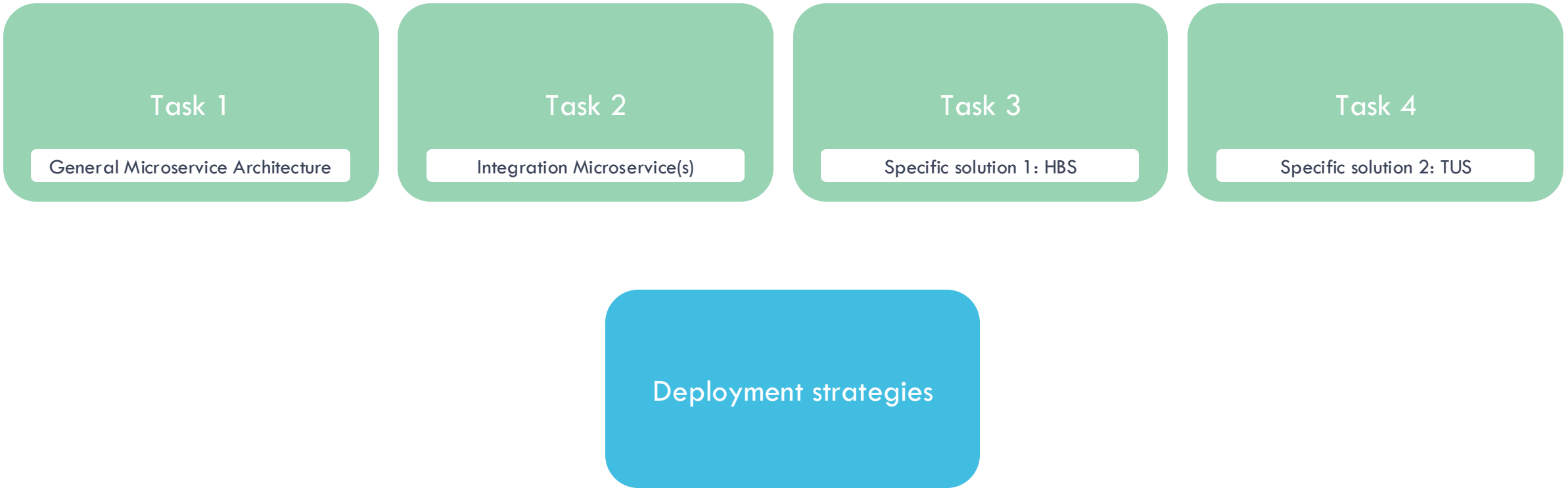
Related stop clusters (home-bank)



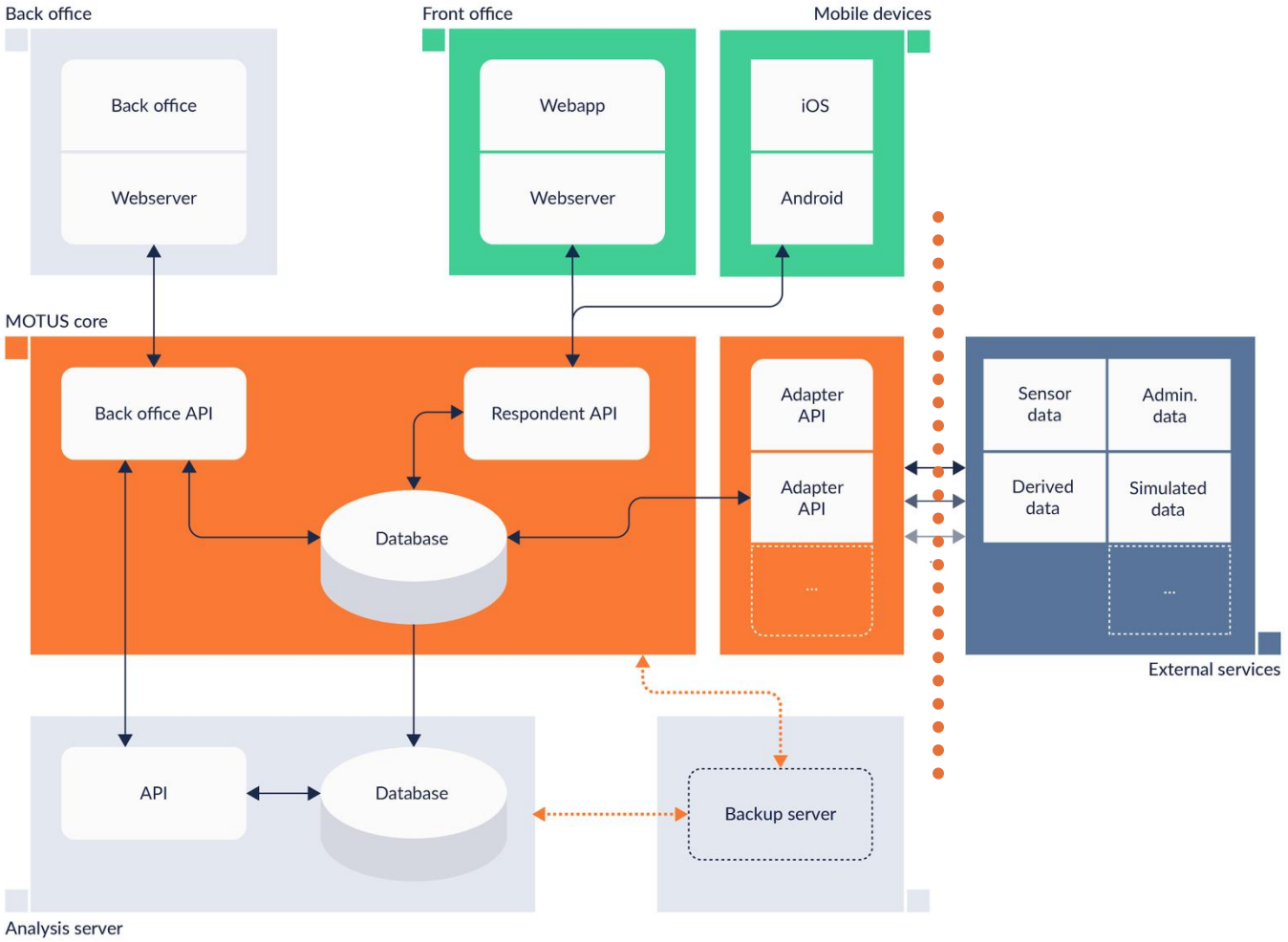
Track (home-bank) + info



Overall task: Deployment strategies

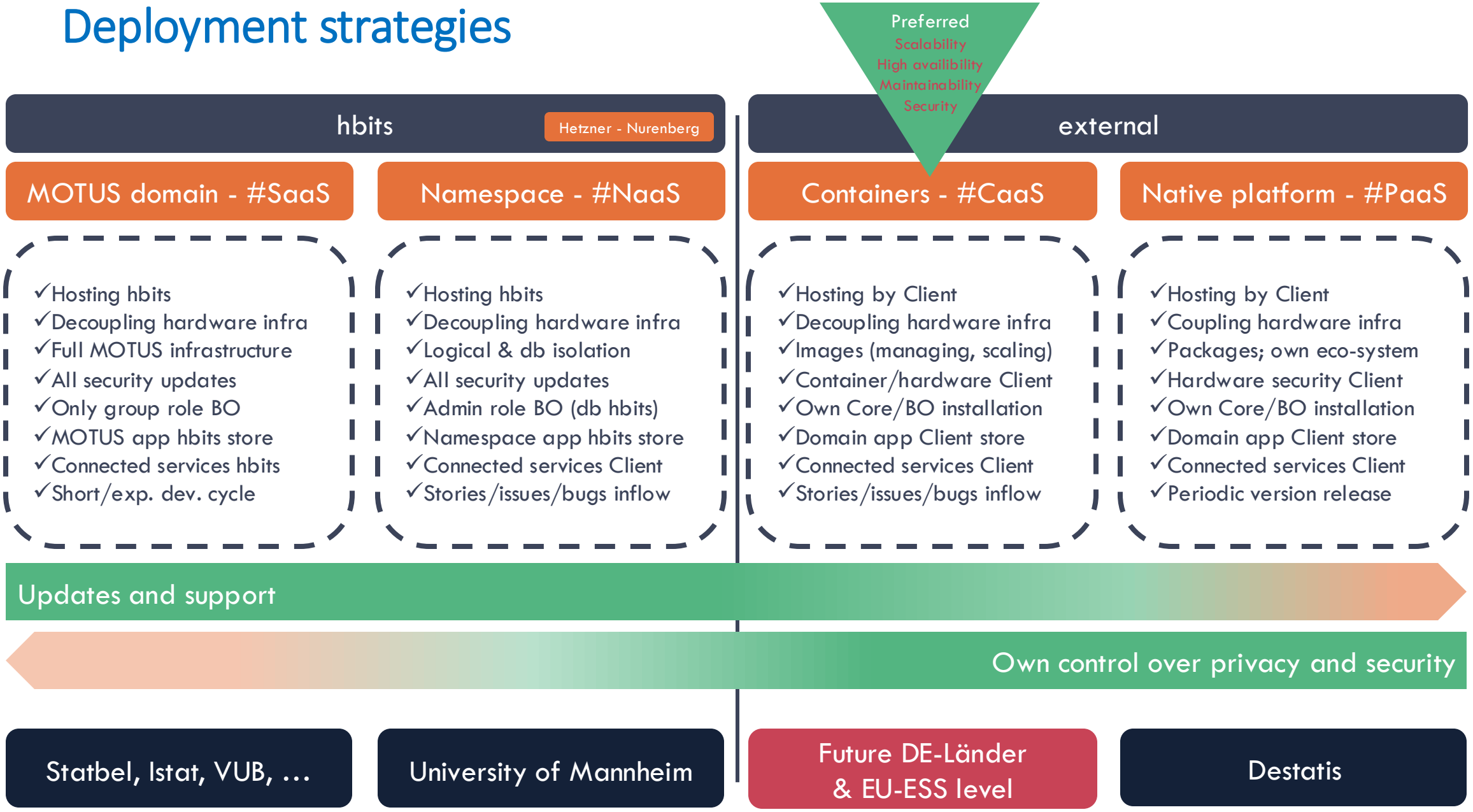


Platform architecture + Microservices = all (micro)services



External to MOTUS

Deployment strategies



Contact details



hbits CV
Joeri Minnen – Managing Director
Joeri.Minnen@hbits.io
Joeri.Minnen@vub.be (as volunteer)