Machine Learning in building stock analysis

BuiltHub Pioneer user training

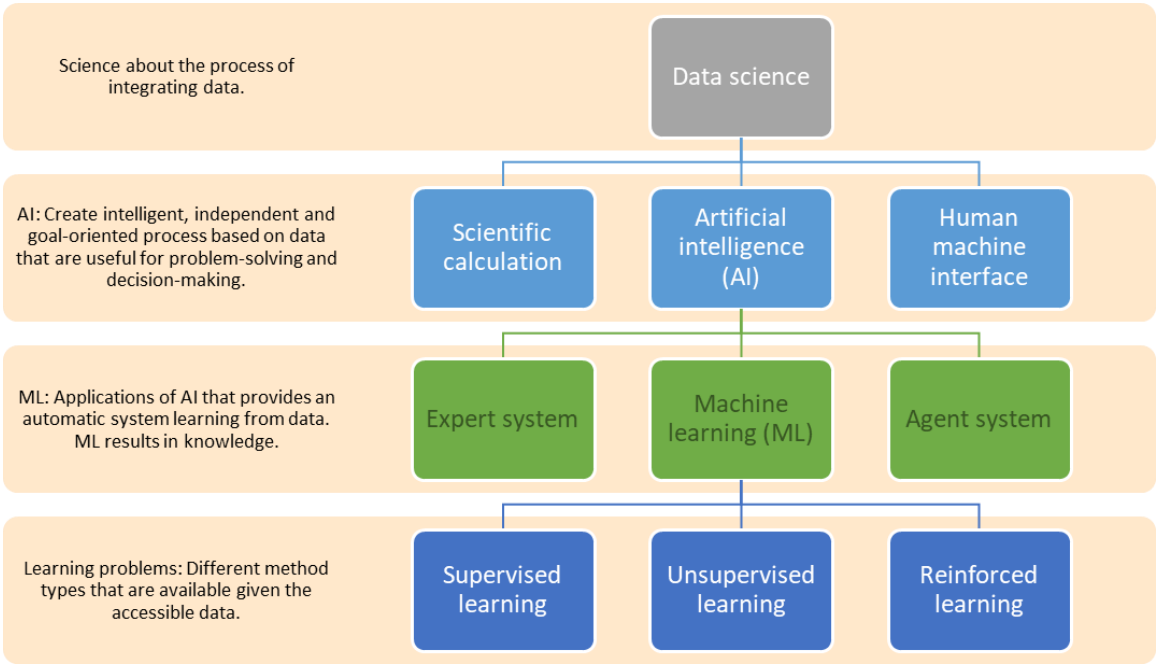
Tim Johansson and Mikael Mangold, RISE

This project has received funding from the EU's Horizon 2020 program under grant agreement no 957026.

Agenda

- Data transformation and clustering
- The application of machine learning
- Matching and merging of building stock information
- 4 Applications
  - Prediction of building types and energy efficiency potentials
    - known features that are not recorded
  - Prediction of cultural heritage values
    - image analysis, known features that are not recorded
  - Prediction of hazardous materials
    - unknown features
  - Usage of fuzzy logic to ascertain building owners
    - data quality, repair of features

Data transformation and clustering



The application of machine learning

Business environment

- Business strategies
- Processes, procedures, systems
- Regulations and market rules
- Market and customer developments and trends

Data environment

Internal data

External data, e.g., open data

Tidy data

Human environment

- Problem formulation
- Accuracy requirement

Decision maker

- Model development

AI engineer

Algorithm environment

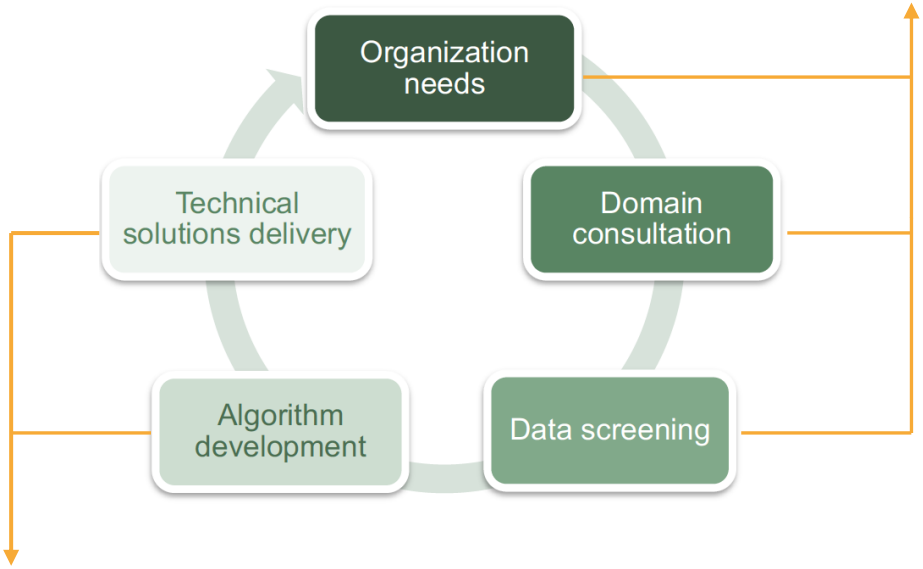
- Model representation
- Optimization
- Software

Value environment

- Knowledge
- Automation
- Decision support

Developing a applied machine learning loop

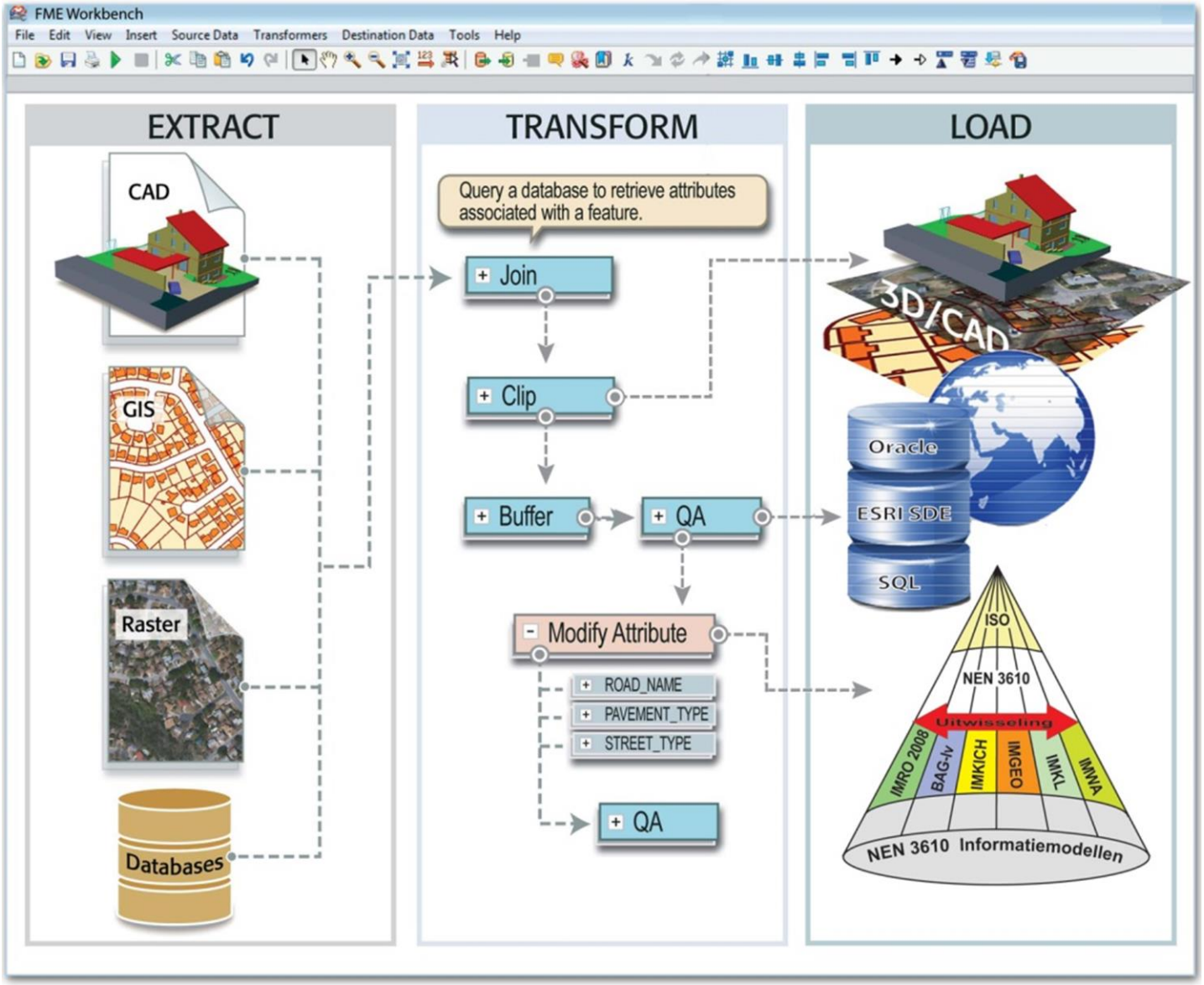
[Predicting the presence of hazardous materials in buildings]



[Enriching the building database for energy retrofitting]

Matching and merging of building stock information

- Matching and merging building stock data is crucial to understand building performance and optimize energy efficiency while minimizing environmental impact.
- However, fragmented data sources, lack of standardization, data quality, privacy concerns, and limited access to data pose challenges to effective data analysis and utilization.
- Extract transform load (ETL) tools can help overcome these challenges and enable efficient data integration and transformation for better analysis and decision-making


FME Workbench
File Edit View Insert Source Data Transformers Destination Data Tools Help
EXTRACT
TRANSFORM
LOAD
CAD
GIS
Raster
Databases
Query a database to retrieve attributes associated with a feature.
Join
Clip
Buffer
QA
Modify Attribute
ROAD_NAME
PAVEMENT_TYPE
STREET_TYPE
QA
3D/CAD
Oracle
ESRI SDE
SQL
ISO
NEN 3610
Uitwisseling
IMRO 2008
BAG-v
IMKICH
IMGEO
IMKL
IMNA
NEN 3610 Informatiemodellen

Extract

1. *Data identified and extracted from various source iteratively*
2. *Analyzed by quality, completeness and fitness for the performance model*
3. *Small datasets used to test ETL workbench with staged data on micro scale before fullscale*

Type of data	Description	Source
Property register	Building data and property information	Lantmäteriet
Property map	Building footprints	Lantmäteriet
Height data	Geotiff and Lidar data	Lantmäteriet
Aerial photo	Mostly Geotiff. Photographs taken from elevated/direct down position	Lantmäteriet
Demographic zones	Squares and borders of municipalities and counties	SCB
Municipality and real estate groups	Databases with municipality categories and real estate groups of the public owned category	SCB, Boverket and SABO
Mining data	Areas affected by mining subsidence, so buildings must be relocated	LKAB
Building data	Various building data delivered from Gällivare and Kiruna municipalities as polygons and points	Gällivare and Kiruna municipalities
Energy data	Energy performance declarations of the multifamily building stock in Sweden	Boverket
Transport data	Transport distances and modes	Google API
Population data	Small test to extract population data by address from webpages	Hitta.se , Eniro.se and upplysning.se
Social data	Data acquired from surveys in Kiruna and Gällivare	Arbetsvetenskap
Energy data	Properties' energy use of district heating in Kiruna, in kWh per day.	TVAB
Building information model	Building information model used for analysis of embedded energy	BIM developed by authors of Paper 3 from drawings

1

Dataset	Multi-family buildings	$A_{BOA+LOA}$ (10^6 m^2)	A_{temp} (10^6 m^2)	Y_c (%)	Y_r (%)	Y_v (%)
Building register	152 470	79.9	-	96.5	39.1	98.2
Property register	-	170.3		96.1	43.7	99.8
EPC register	138 444	-	211.7	100.0	-	-
Energy Agency*	-	191.6	-	-	-	-

*Swedish Energy Agency report 2014

2

3

Transform

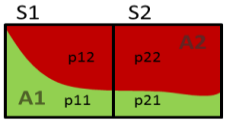
Definition of performance and evaluation criteria

Transformation
Create new attributes
Divide attributes
Modify attributes

Transformation
Expression evaluation
Aggregation
Formatting
Sorting
Duplicate remover

2

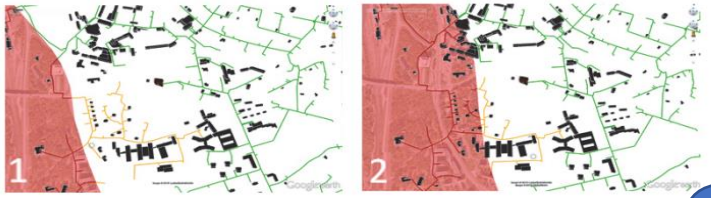
Attribute transformation



Square	Shape	Population
S1	Polygon	112
S2	Polygon	250

Zone	Shape	Population
A1	Polygon	p11*112+p21*250
A2	Polygon	p12*112+p22*250

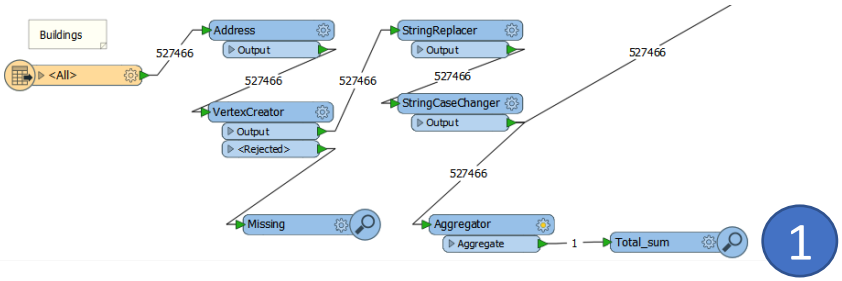
p11...p22 are parts of the areas S1 and S2 in %



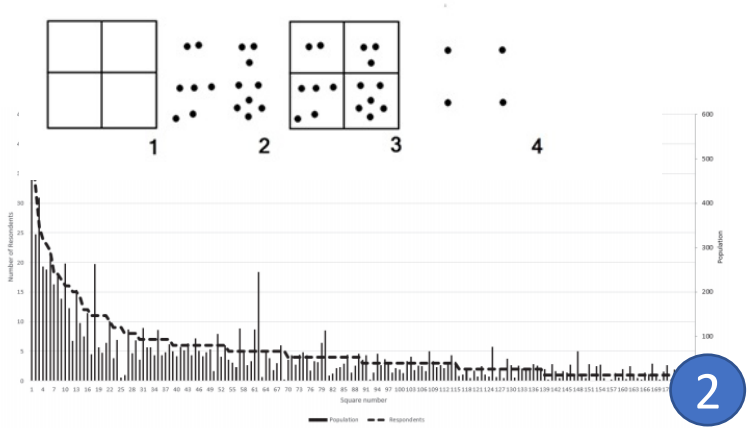
3

Spatial transformations

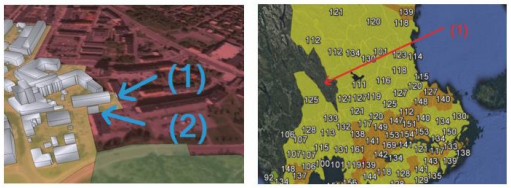
Load

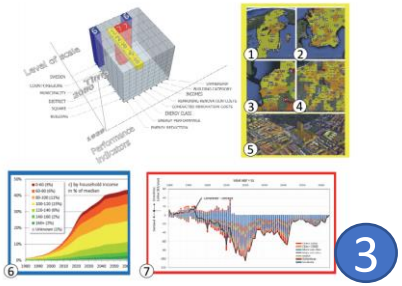


Process validation



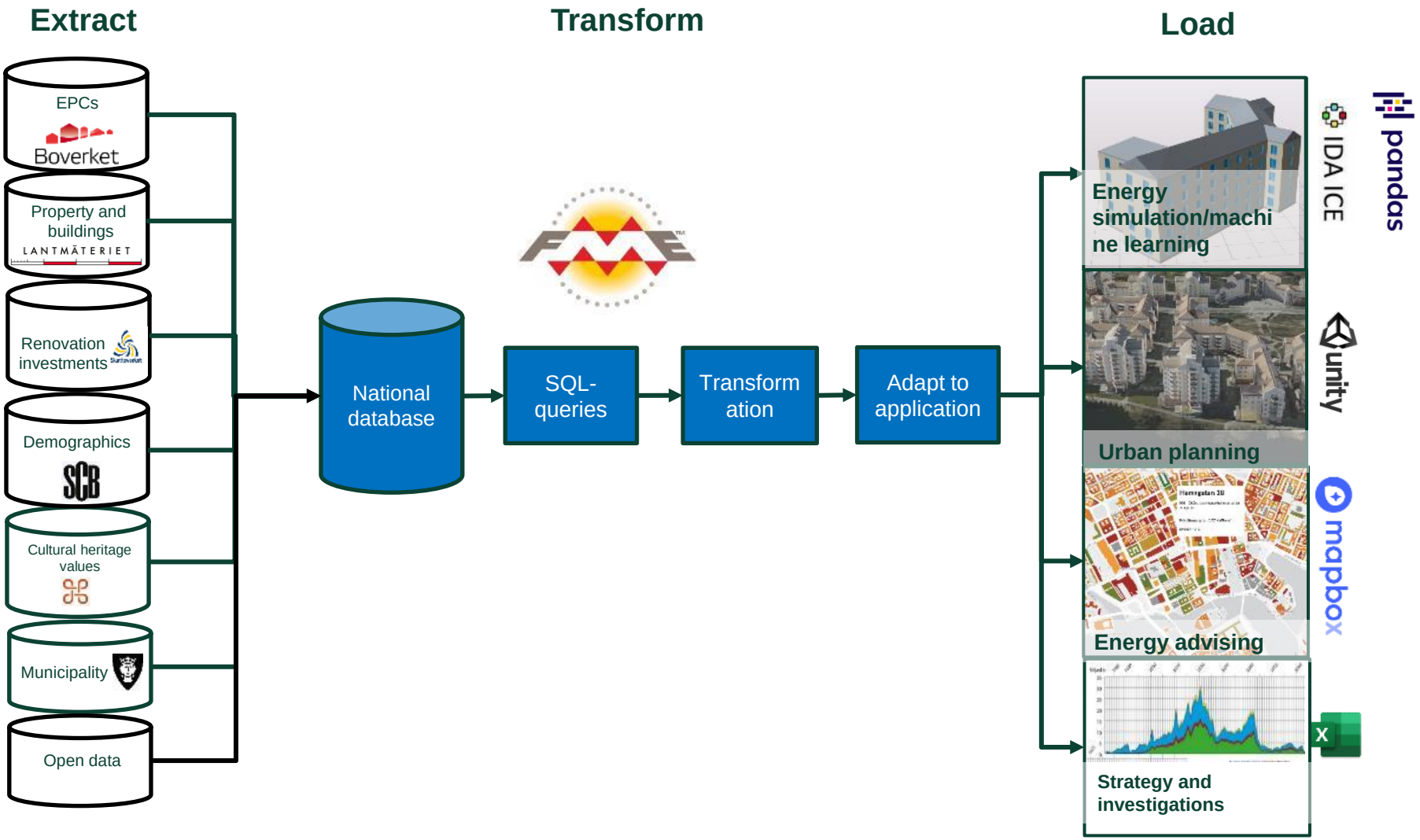
Spatial validation





Spatial visualization and analysis

Matching and merging of building stock information



Application 1
Prediction of building types and energy efficiency potentials

Background

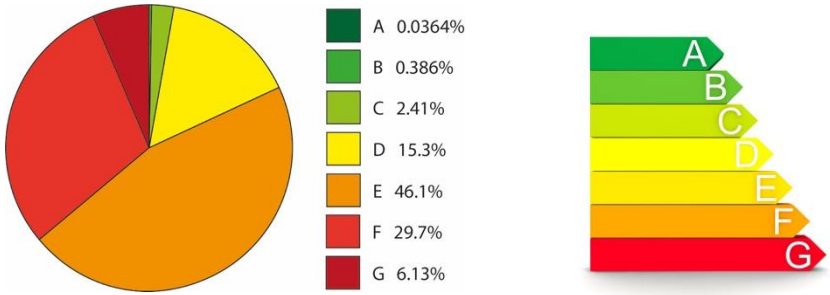
- A majority of the Swedish multifamily building stock was constructed 1945-1975
- The building stock consists of tower blocks, slab blocks (before 1960), slab blocks (after 1960) and panel blocks
- Significant needs for refurbishment
- Opportunity to incorporate energy conservation measures when refurbishing

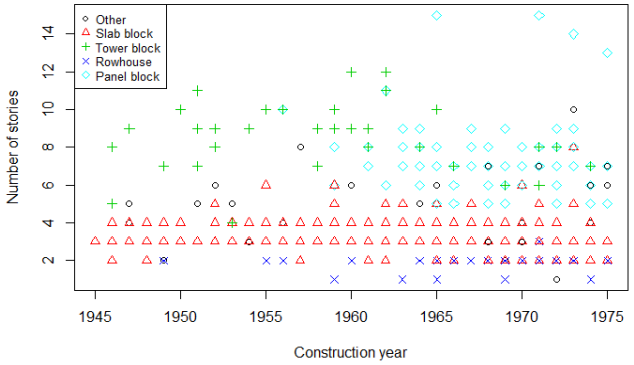




Method – enrich data

- Information from EPCs (energy use, heated area, construction year, address, no of storeys, stairwells, apartments etc.)
- Building characteristics were collected from 514 ocular Google Street View observations
- Machine learning algorithms were then used to find patterns correlating information from EPCs with characteristics from Google Street View observations

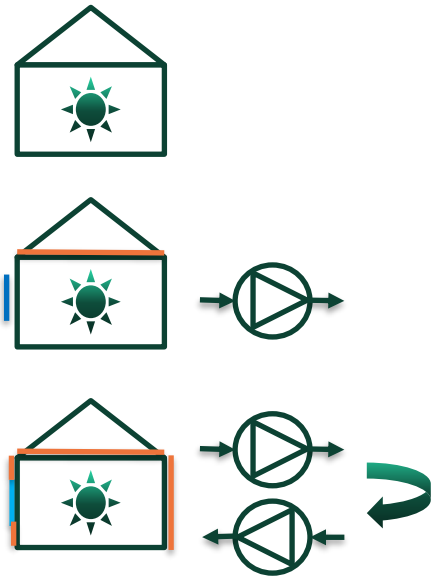




Energy efficiency potential

3 different renovation packages were developed

1. Measures to optimise the operation of the building
2. Package 1 + Change to more effective pumps and fans and put additional insulation in the attic and add extra pane to existing windows
3. Package 1+2+ New ventilation system with heat exchange from exhaust air, new windows, and 10 cm additional insulation at the building envelope



Energy efficiency potential

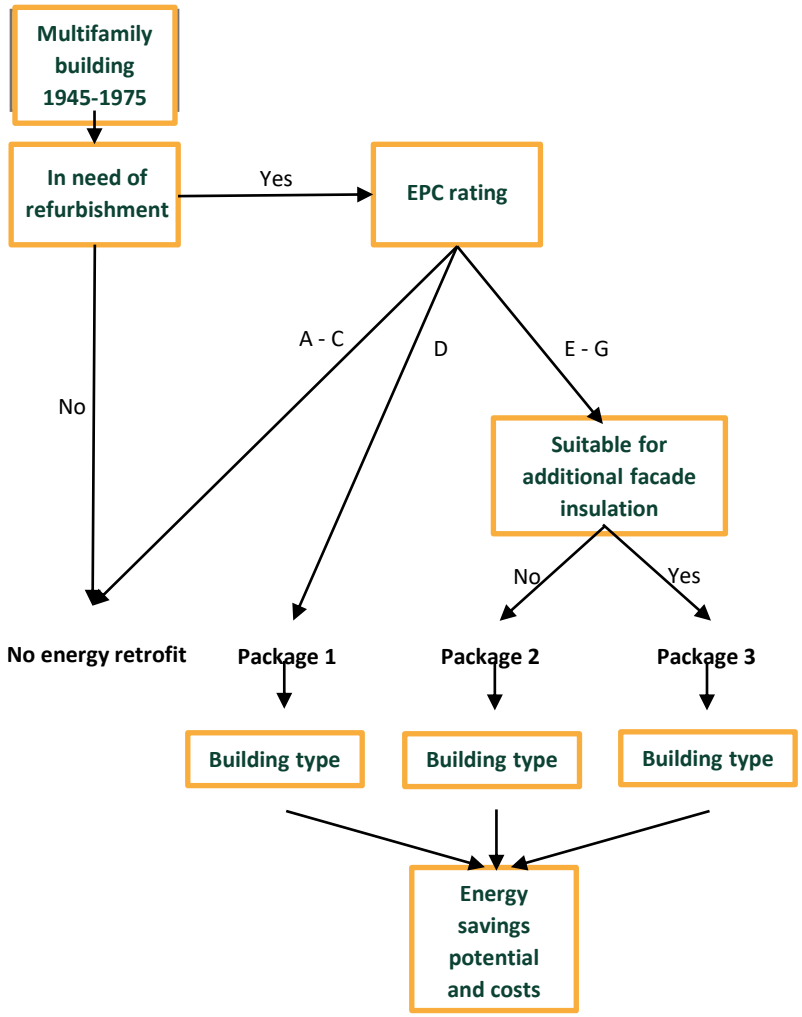
- Cultural historical values and specific building features must be preserved such as brick façade.
- No eaves overhang involves a need for extensive inventions adjusting the existing roof to a thicker façade.
- Thus, two characteristics; façade material and eaves overhang, determine whether additional façade insulation (package 3) is suitable or not.

Not suitable

Suitable

Method – tailored renovation packages

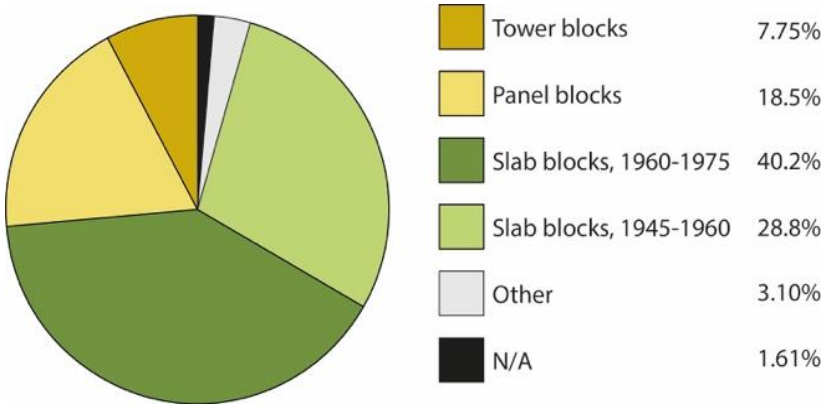
- Based on the EPC rating a tailored renovation package was allocated to each specific building in the building stock.
- The energy savings and cost for renovation measures was estimated.
- Limited to the Swedish multifamily building stock constructed between 1945 and 1975



Results

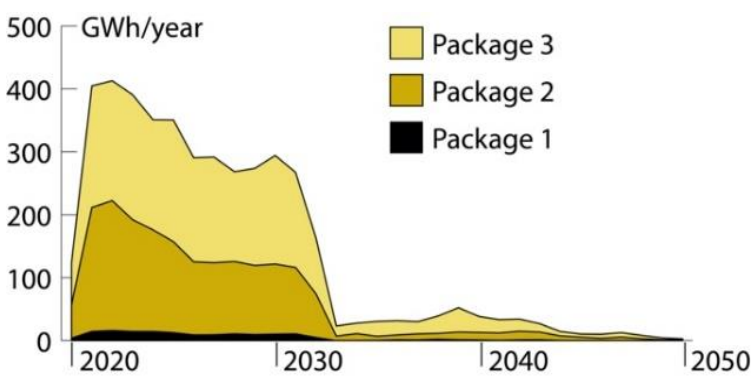
- With the chosen model SVM (support vector machines) building type is predicted with an accuracy of 88,9 %

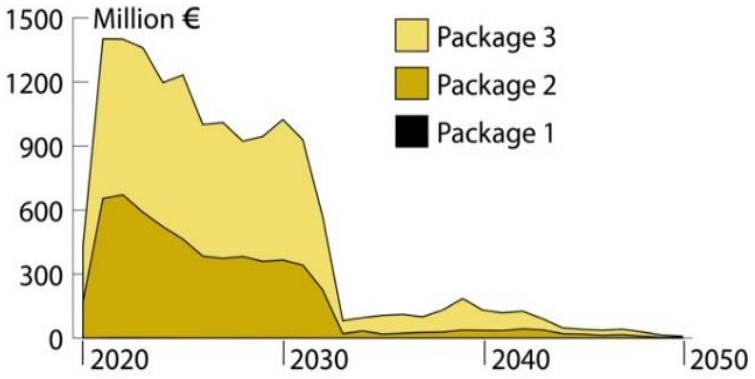
- With a model accuracy of 72.5%, it was predicted that 32.0% of all multifamily buildings are suitable for additional façade insulation (energy retrofitting package 3)



Building Type	Eaves Overhang and not Brick Façade [%]
Slab blocks, <1960	63.9
Slab blocks, 1960–1975	22.0
Panel blocks	6.81
Tower blocks	26.4
All building types in multifamily building stock 1945–1975	32.0

Final results





Figures: Energy savings potential from the different energy retrofitting packages and the associated costs.

Application 2

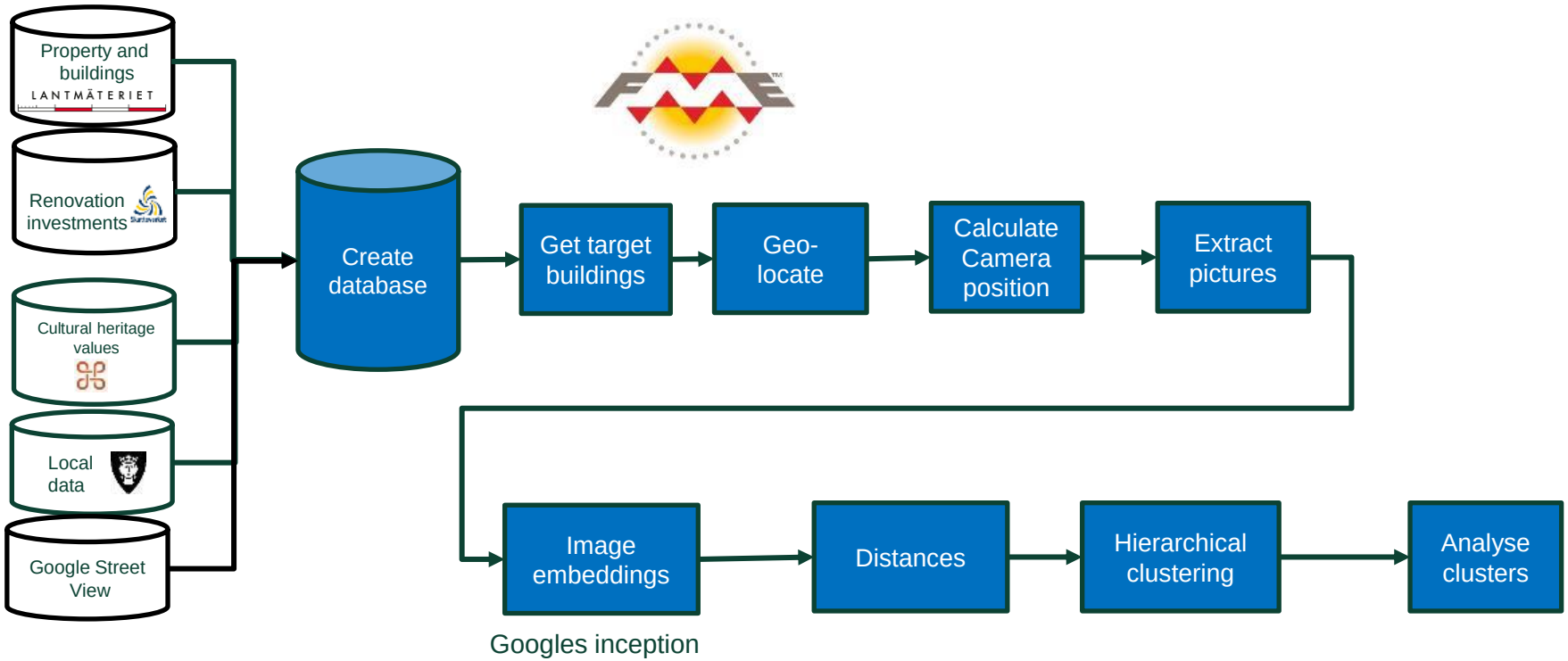
- Prediction of cultural heritage values

- Currently, there is no complete building registry that includes cultural heritage values, which are scattered across municipal and regional databases and plans.
- In Sweden, cultural heritage values are often missed due to buildings not being identified, which means that stakeholders are not aware of the cultural heritage values.
- There is an important task to classify buildings according to their cultural heritage values, and with the help of machine learning, inventory data from different parts of Sweden can be scaled up to a national level.

Extract

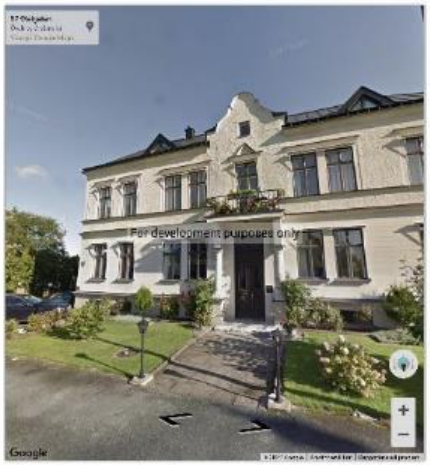
Transform

Load

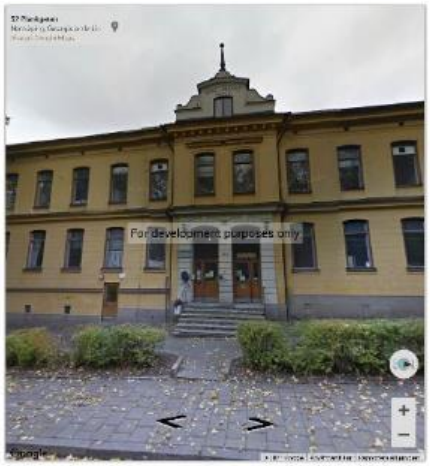


1570.html

1577.html



3.html



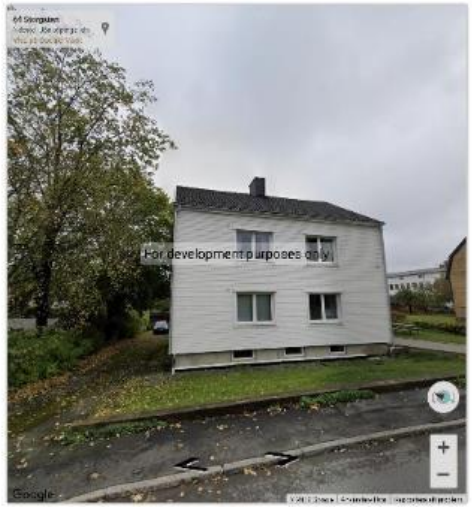
455.html

1219.html

332.html

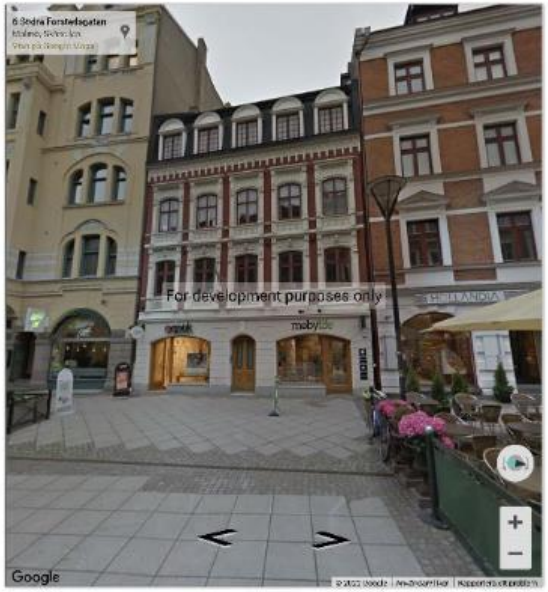
577.html

883.html

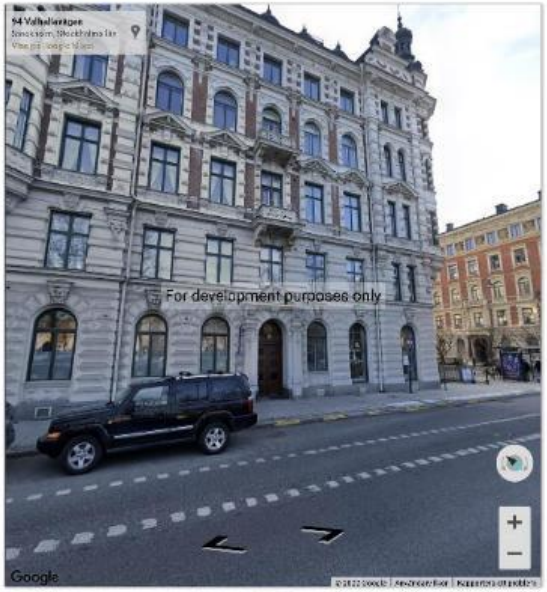


918.html

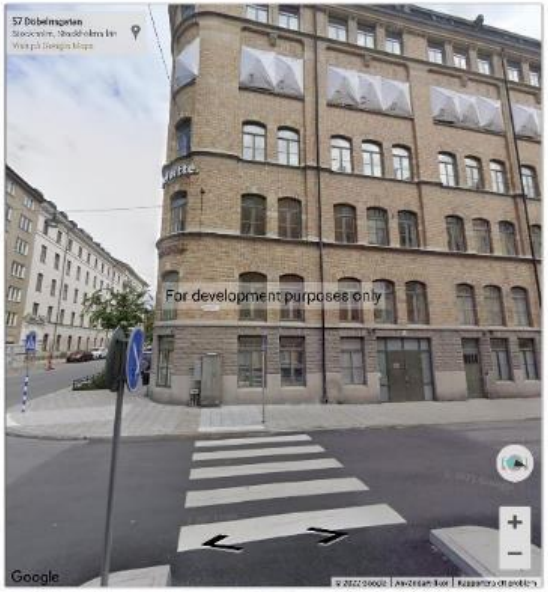
BuiltHub



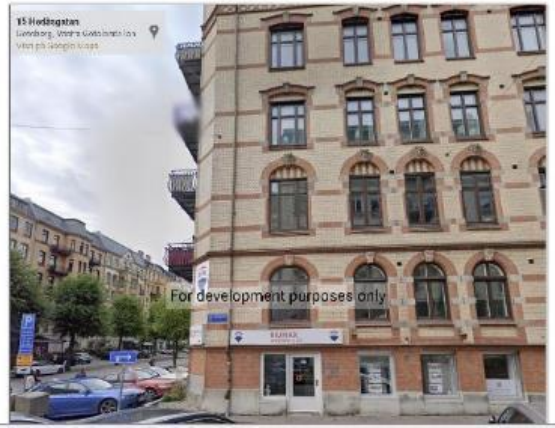
1254.html



261.html



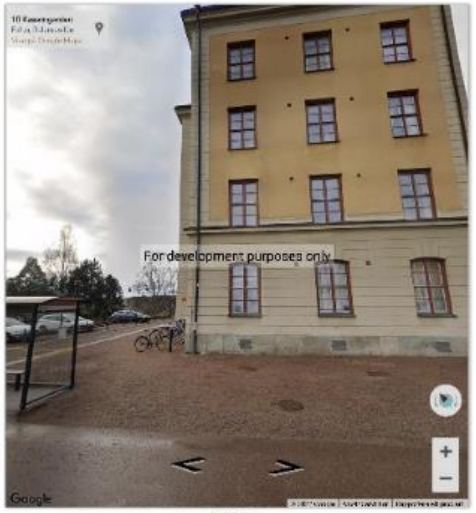
360.html



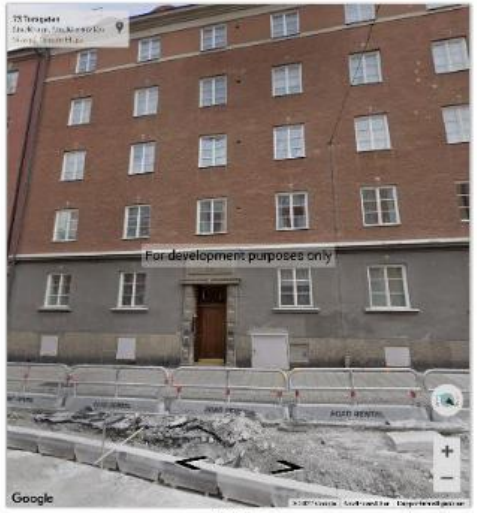
BuiltHub

1234.html

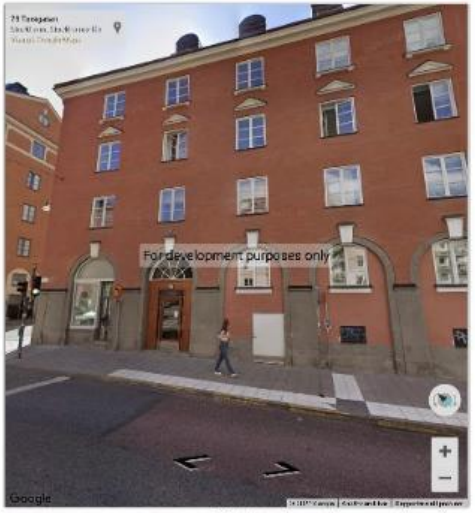
46.html



693.html



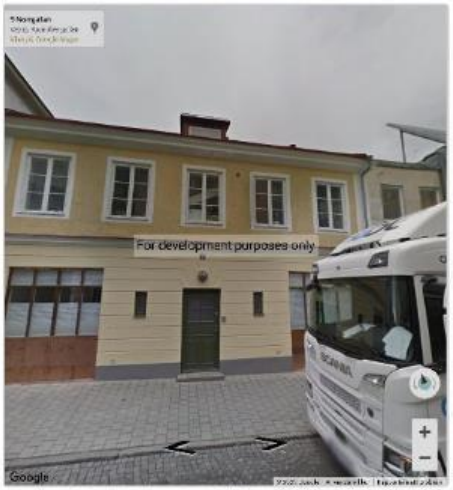
705.html



964.html

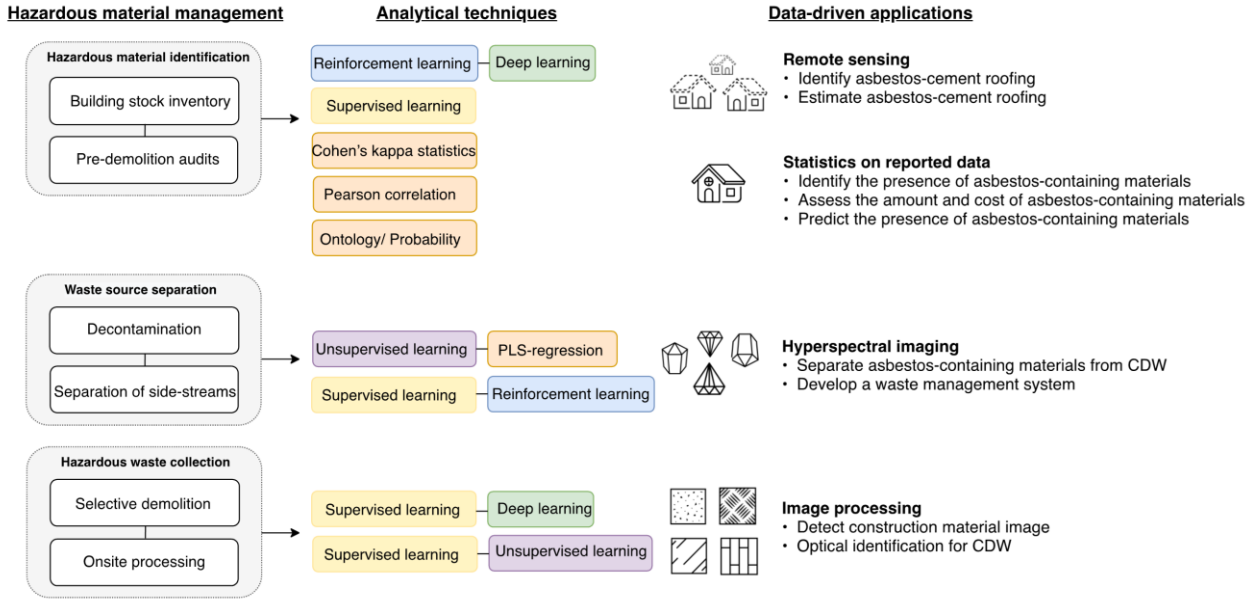
1287.html

1415.html

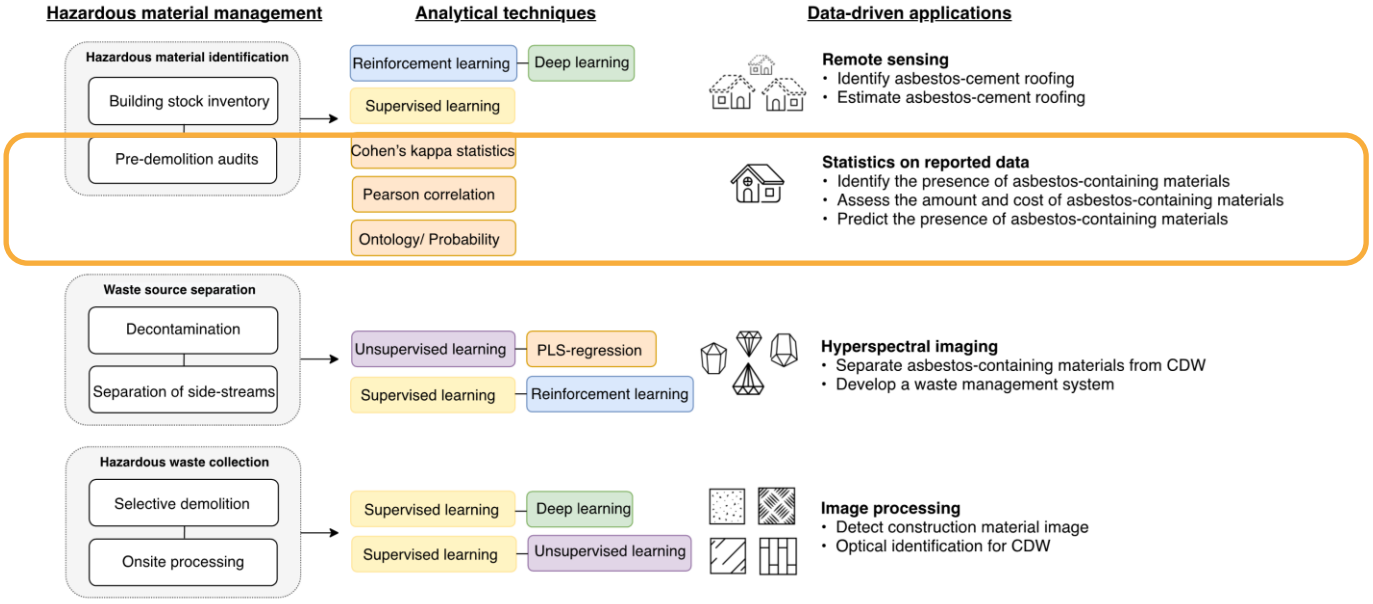


1627.html

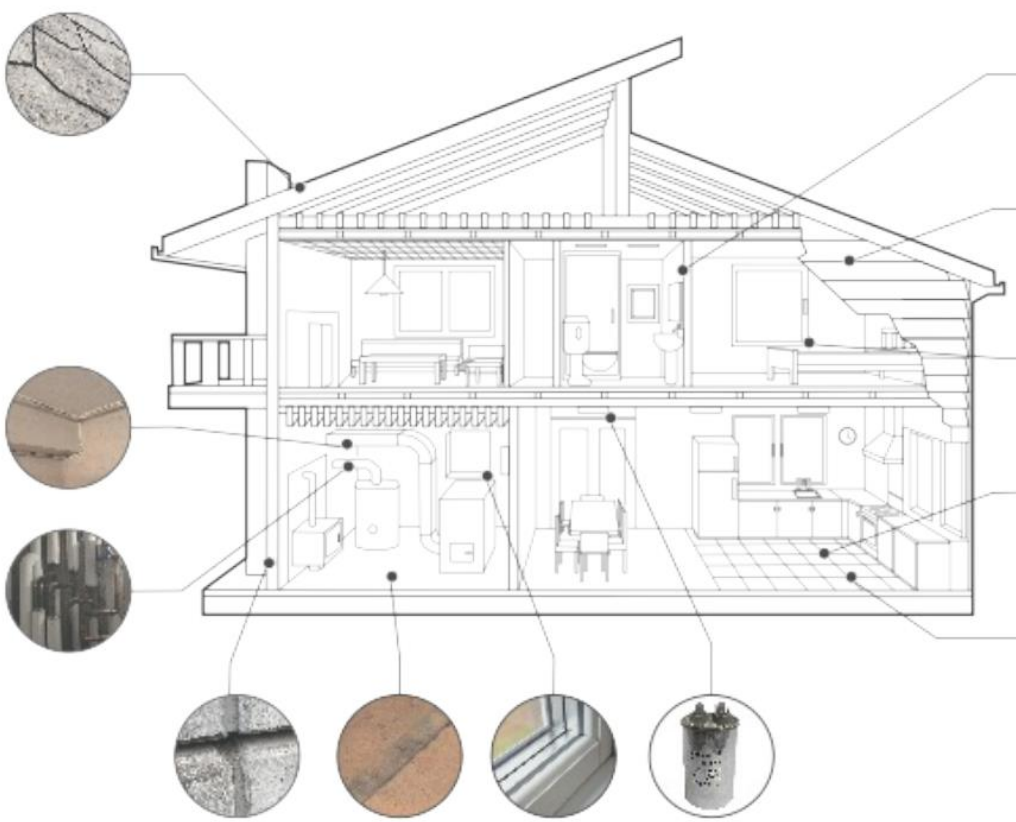
Application 3
– predicting the presence of hazardous waste in the building stock



Application 3
– predicting the presence of hazardous waste in the building stock

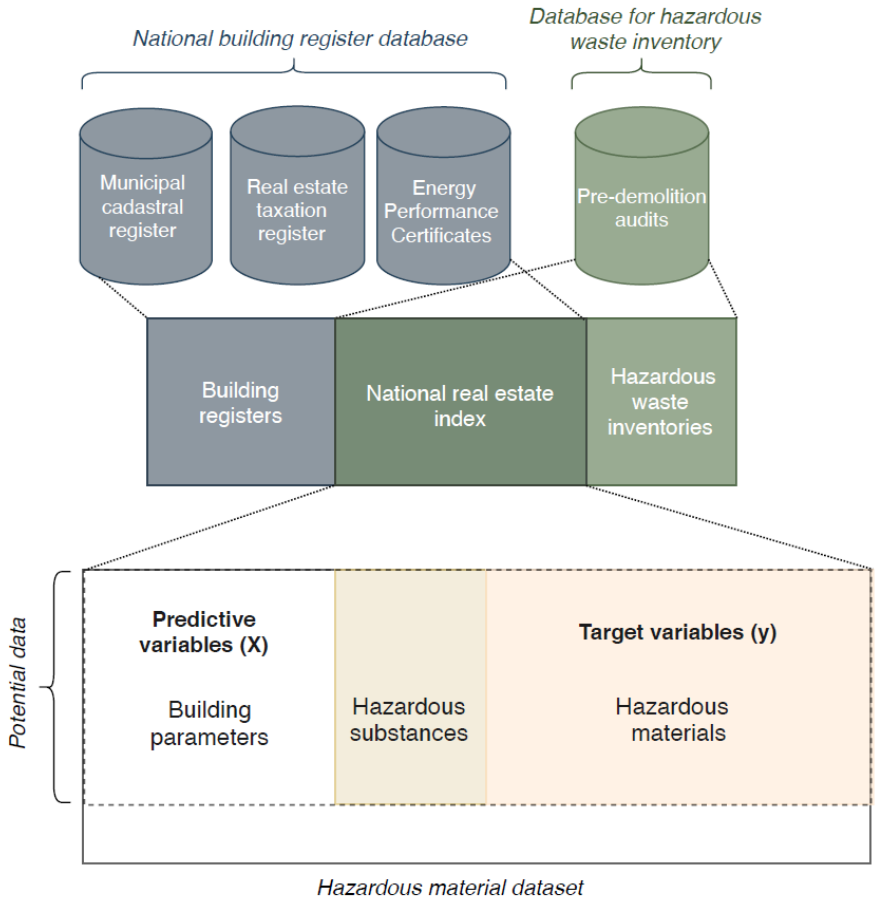


Pre-demolition audits quantify hazardous materials in buildings



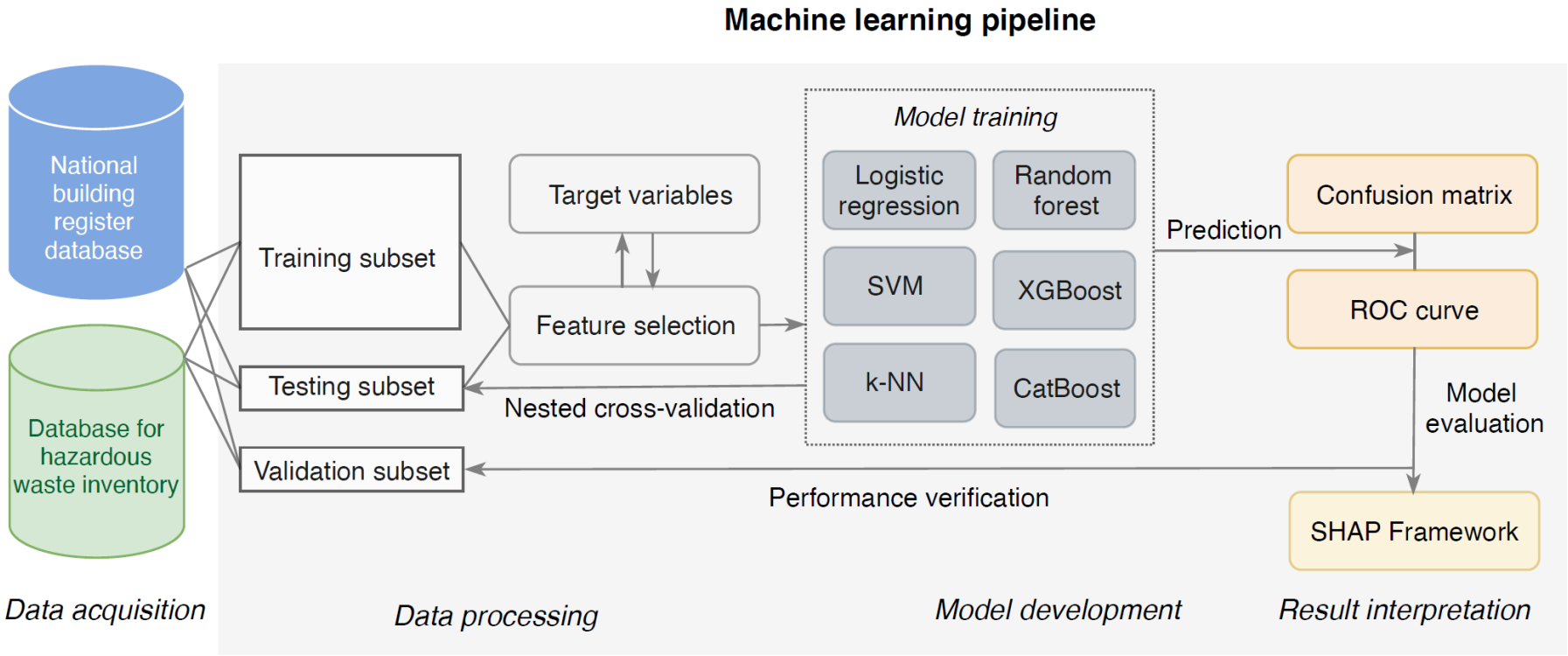
* PCB (Polychlorinated biphenyls)

Workflow for creating and validating a hazardous material dataset



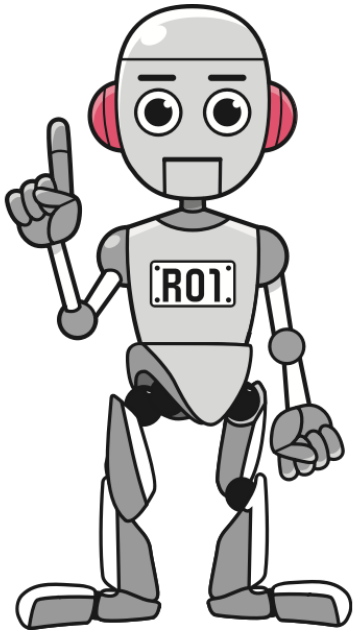
1. Data collection
2. Data matching
3. Data validation

Development of a machine learning pipeline for predictions



Application 4

- Usage of fuzzy logic to ascertain building owners



2016 2021

Din Bostad FK AB = Heimstaden FK AB?

New owner?

Problem

1. Comparing changes in property ownership over time can be difficult due to the old organization numbers remains unchanged and hierarchical company structures.
2. Small changes in organization names can further complicate the identification of ownership changes.
3. By utilizing this information, it is possible to study the impact of rental housing acquisitions on the largest rental developers.
4. Understanding the transitions in ownership can help us to understand the transformation of the building stock and support decision making

Method

1. Two company names were compared to identify any differences in spelling or syntax using a fuzzy matching algorithm that generated a score based on the level of similarity between the names.
2. If the company names were an exact match, or if the fuzzy score was very high, this suggested that no change in building ownership had taken place.
3. If there were slight variations in the names, or if the fuzzy score was lower, further investigation was conducted to determine whether a change in building ownership had indeed occurred, despite the retention of the same identification number.

Example:

Company name 2016/2021		Score	Org number 2016/2021	
BRF HÄSÄNGEN	Riksbyggen Bostadsrättsförening Häsängen	0,69	776400-0241	776400-0241
Huge Fastigheter AB	Huge Bostäder AB	0,69	556233-5900	556149-8121
Huge Fastigheter AB	Huge Bostäder AB	0,69	556233-5900	556149-8121
Sunne Bostads Aktiebolag	Sunne Fastighets AB	0,69	556042-8921	556042-8921
Ulricehamns Förvaltning AB	Bogesund Förvaltning AB	0,69	556954-5717	556966-5903
Ulricehamns Förvaltning AB	Bogesund Förvaltning AB	0,69	556954-5717	556966-5903
KRAMBO BOSTADS AKTIEBOLAG	Krambo Aktiebolag	0,69	556345-8701	559200-9004
Huge Fastigheter AB	Huge Bostäder AB	0,69	556233-5900	556149-8121
KRAMBO BOSTADS AKTIEBOLAG	Krambo Aktiebolag	0,69	556345-8701	559200-9004
Ulricehamns Förvaltning AB	Bogesund Förvaltning AB	0,69	556954-5717	556966-5903
Din Bostad FK AB	Heimstaden FK AB	0,69	556712-8953	556712-8953
KRAMBO BOSTADS AKTIEBOLAG	Krambo Aktiebolag	0,69	556345-8701	559200-9004
BRF STAMGÅRDSPARKEN	HSB Bostadsrättsförening Stamgårdsparken i Sundbyberg	0,69	769607-9743	769607-9743
Riksbyggen Bostadsrättsförening Göteborgshus nr 8	Bostadsrättsföreningen Göteborgshus 8	0,69	757201-7643	757201-7643
KRAMBO BOSTADS AKTIEBOLAG	Krambo Aktiebolag	0,69	556345-8701	559200-9004

Thank you!


BPIE


CARTIF


everis
an NTT DATA Company


eurac
research


ICLEI
Local Governments
for Sustainability
EUROPE


RI.
SE


TEAM
sympraxis


TU
WIEN
TECHNISCHE
UNIVERSITÄT
WIEN
Vienna | Austria

BuiltHub

This project has received funding from the EU's Horizon 2020 program under grant agreement no 957026.