

# An LLM-powered Social Digital Twinning Platform

RED Seminar Series



Önder GÜRCAN



27.03.2025



# Understanding the Complexity of Social Issues through Simulations



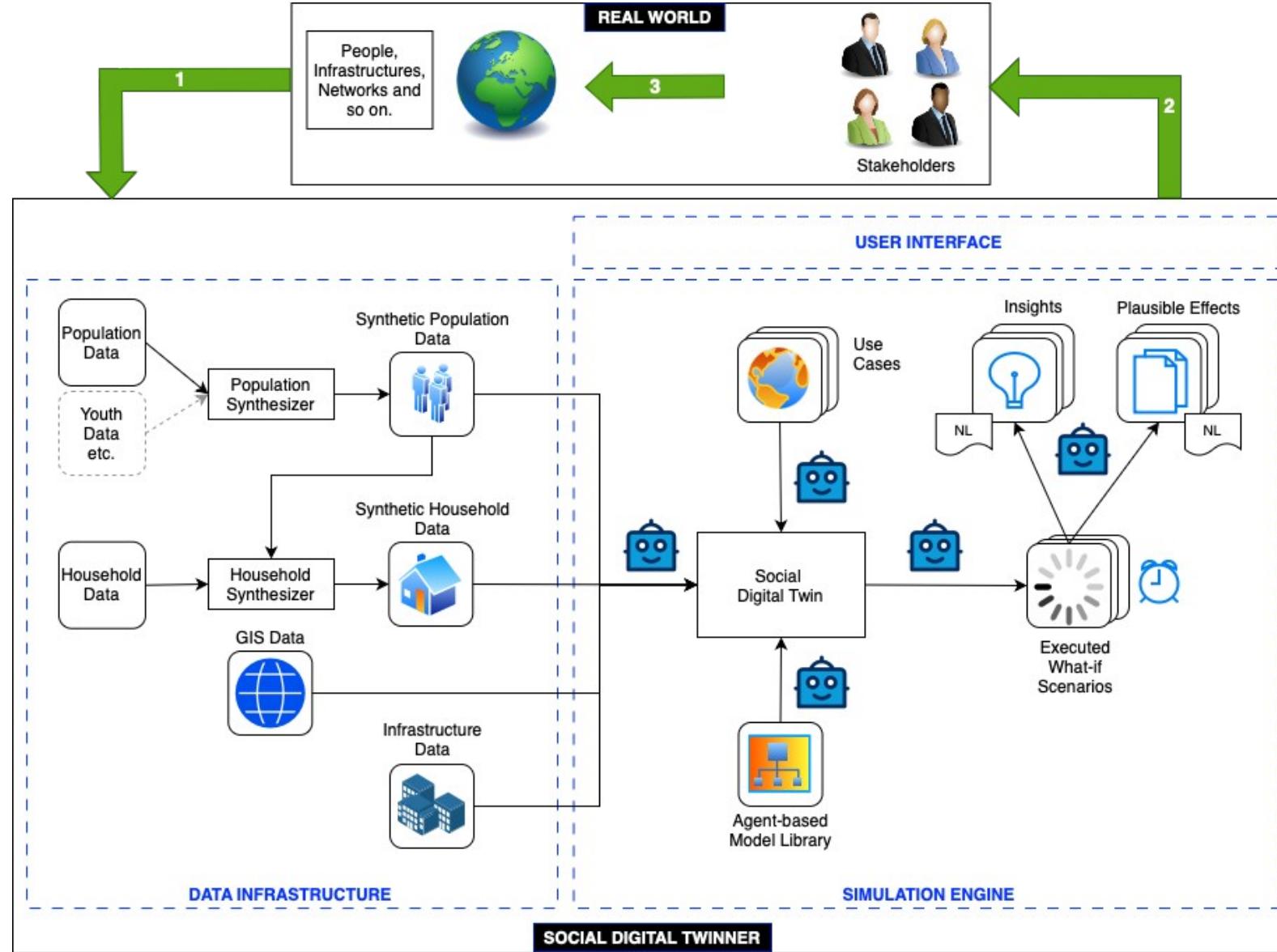
- Social issues are **complex**.
  - Numerous interacting variables,
    - including human behaviour,
    - cultural norms, economic factors,
    - and political dynamics
  - Constantly changing over time
- Traditional studies **cannot capture**
  - dynamic, non-linear interactions
  - and emergent phenomena.
- **Advantages of simulations**
  - Analyse and predict outcomes,
  - Controlled, replicable analysis
- **Social digital twinning:**
  - Collect as much as real-world **data** possible,
  - Recreate real-world **social conditions** through simulation,
  - Recreate people's **actions** and **thoughts** through simulation,
  - Run simulations using **various** (what-if) **scenarios**,
  - Test **policy** changes,
  - Forecast **trends** by calculating indicators, e.g. CO2 emission by area,

# Motivation

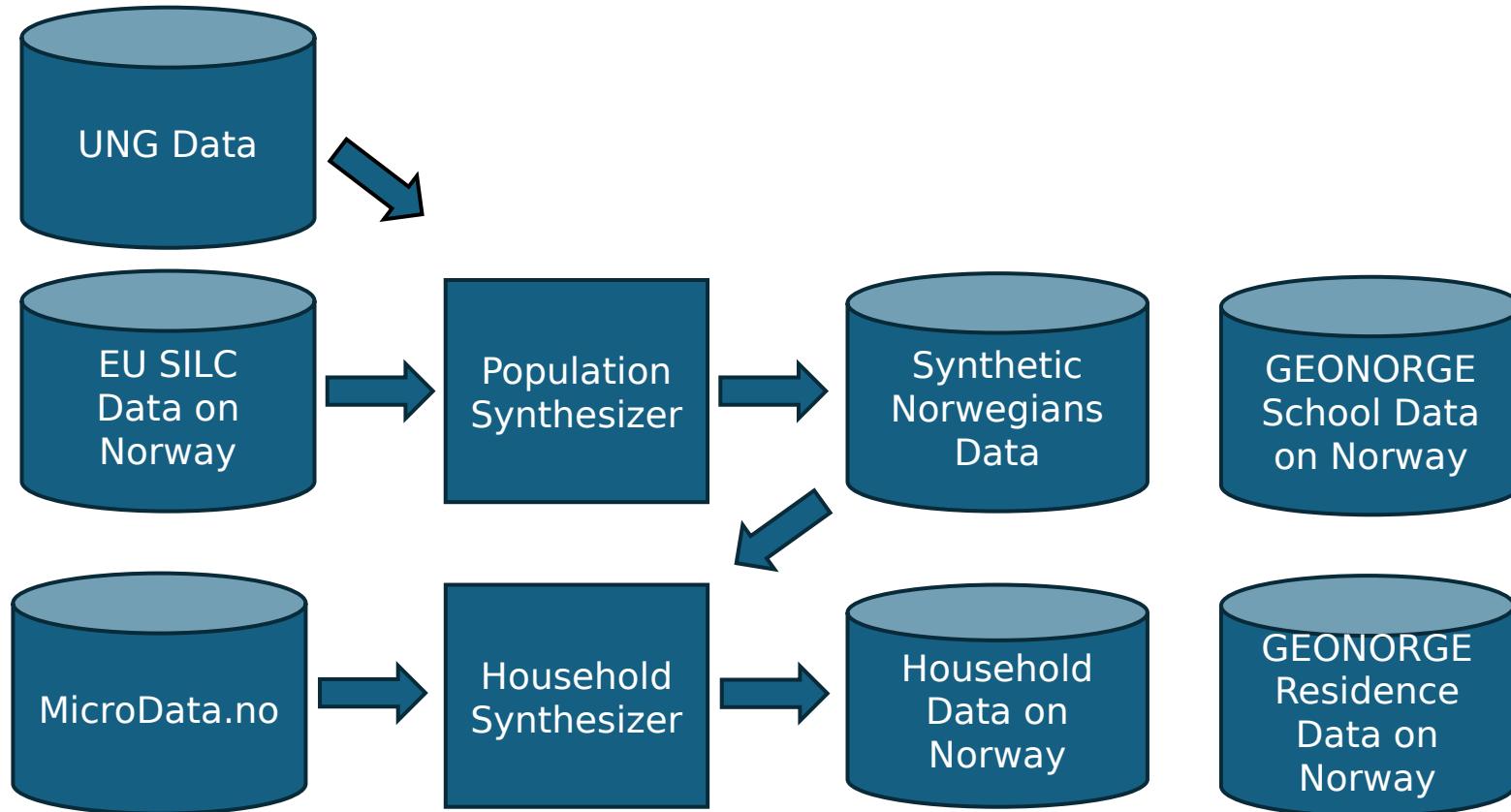


- Most research on SDT applications focuses
  - on the smart city domain.
- Most existing SDT solutions remain tightly focused on specific problems
  - rather than offering a flexible, generic platform.
- Moreover, they often require advanced programming knowledge to set up, customize, or improve,
  - which creates a significant barrier for decision-makers, researchers, or community stakeholders who lack deep technical expertise.

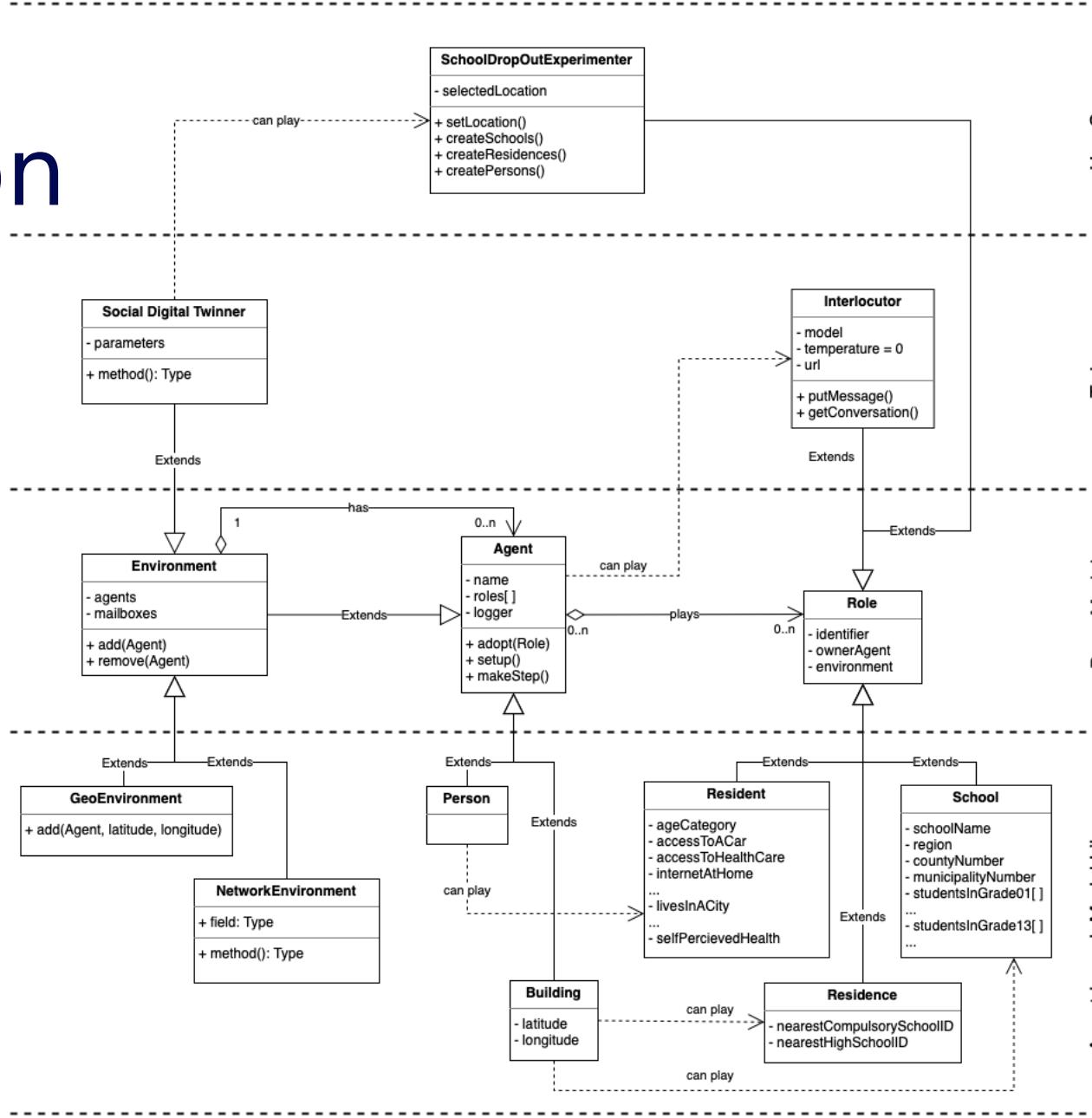
# Social Digital Twinner



# Twinner: Data Infrastructure



# Twinner: Simulation Engine



"Be as concise as possible in your answers. If you do not know the answer to a question, simply answer as 'I do not know.'."

You are {ownerAgent.name}, fulfilling the role of {role.name} in {environment}.

You have access to {role\_specific\_data}. You are empowered to act directly as {role.name}.

Analyze the user's question or request, leverage your knowledge and the provided data, and generate the best possible answer or solution.

If you need additional reasoning steps, please outline them clearly.

Then, finalize your response as {role.name}.

# Twinner: User Interface

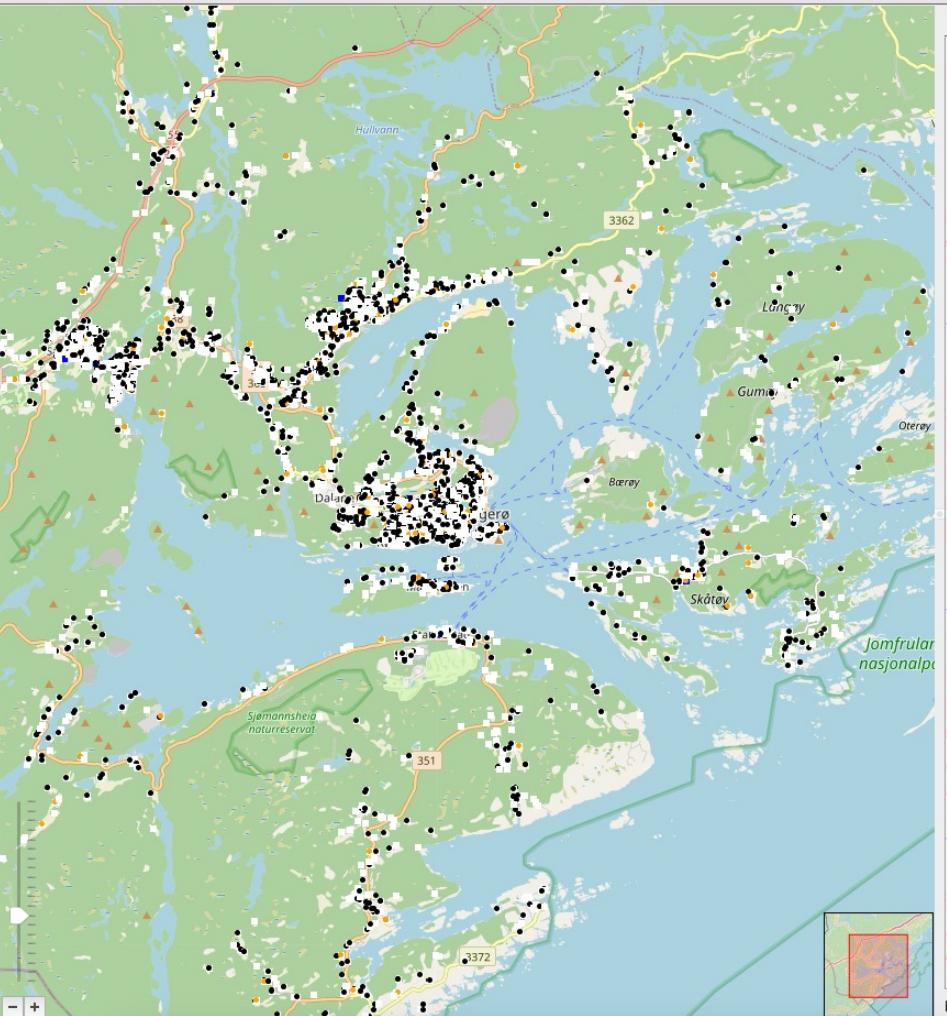


Social Digital Twinner

File View Help

Configure... Setup Run Pause Episode: 0 of 1 Tick: 0

Norway



Social Digital Twinner

Using this scenario, you can try to answer research questions such as:

(1) Why do students drop out of high school?  
(2) How can their situation be improved?

If you want to proceed, I can create synthetic persons, schools and residences for you. What would you like to involve in the scenario?

You: I want to create schools.

Social Digital Twinner: I added compulsory schools (i.e. grunnskole) and high schools (i.e. videregående skole) for you. The point data, addresses, and contact information are sourced from Nasjonalt Skoleregister, managed by Utdanningsdirektoratet, and retrieved via GeoNorge. Information on the number of students per grade at each school is sourced from Grunnskolens Informasjonssystem (GSI) by Utdanningsdirektoratet.

You: I want to create residences.

Social Digital Twinner: I created residences for you using data compiled from publicly available versions of Matrikkelen. Point data and building type are sourced from Matrikkelen Bygningspunkt, while address, residential unit, and apartment level are from Matrikkelen Adresse Leilighetsnivå to ensure that all apartments in a building are included.

You: I want to create persons.

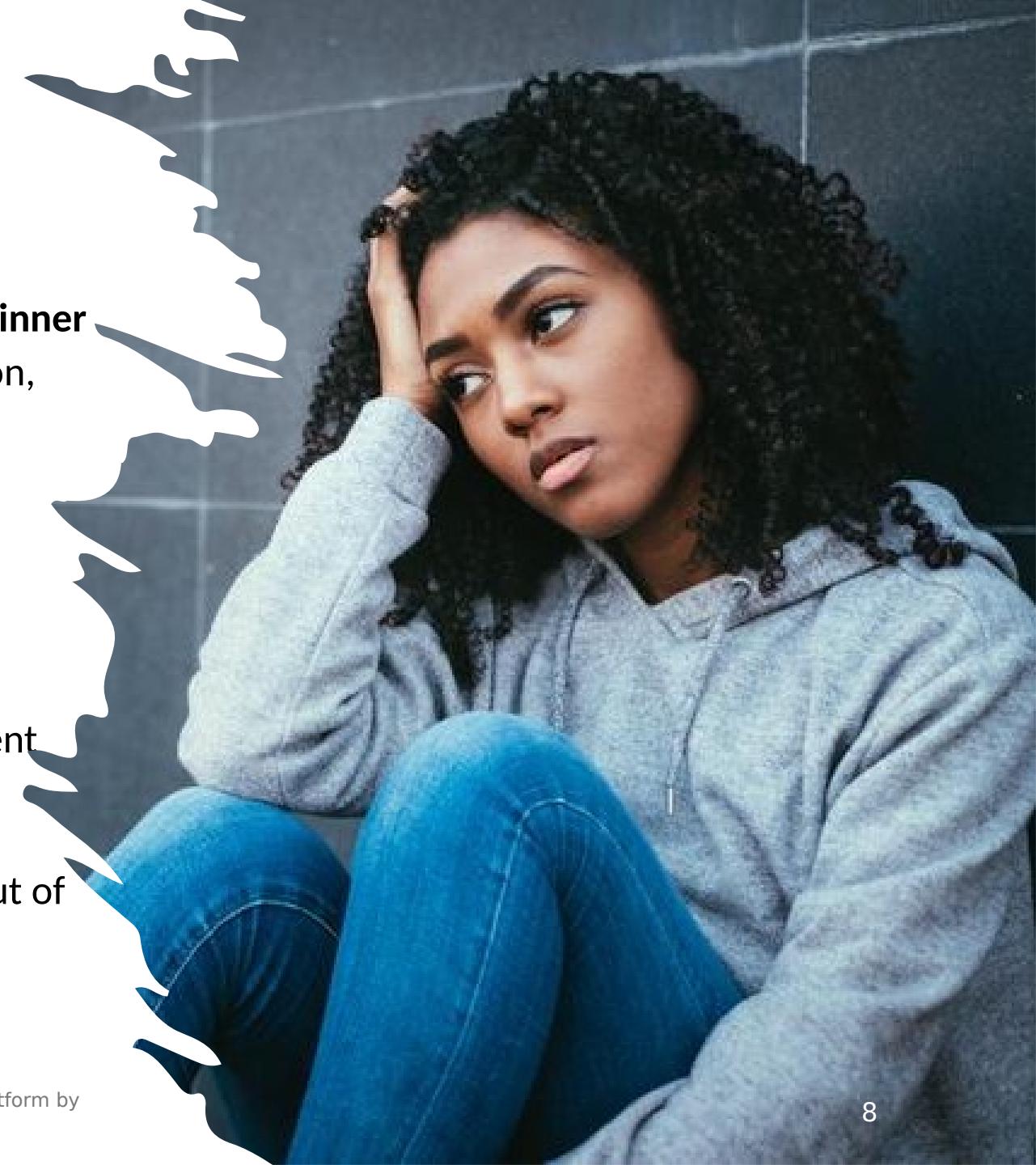
Social Digital Twinner: I created persons for you using a synthetic population generated from a Wasserstein Generative Adversarial Network (WGAN) model trained on balanced EU-SILC 2020 data from Statistics Norway. The population for Norway is created by matching demographic profiles based on age, gender, and

Prompt:

Send

# Digital NEETs in Norway (DN2) Pilot

- DN2 Pilot is the first pilot case study of **Social Digital Twinner**
  - based on the phenomenon of NEET (Not in Education, Employment, or Training).
- In this pilot, we study the school drop out of NEETs.
  - Location: Kragerø, Norway
  - Scenario
    - Students from rural and urban areas attend secondary schools, and high-schools.
    - If there is no high-school in a rural area, a student may attend to a school in closest urban area.
  - Questions
    - Why do students who live in rural areas drop out of high school?
    - How can this situation be improved?



# Demo



- [https://drive.google.com/file/d/1Ju6vZHk7djYxszjHkag6VqWpAufObXn/view?usp=share\\_link](https://drive.google.com/file/d/1Ju6vZHk7djYxszjHkag6VqWpAufObXn/view?usp=share_link)

# Discussion



- Privacy
  - The privacy of personal data was considered when retrieving georeferenced data.
  - The synthetic population is anonymised based on training anonymised and restricted-use data from EU SILC.
  - The synthetic population is allocated randomly to residences and schools.
- Security
  - The data is securely accessed mitigating the risk of unauthorized breaches.
- Democratisation
  - On-demand, interactive social digital twins via natural language interfaces.
    - The platform lowers technical barriers, empowering non-specialists to simulate complex use cases.

# Conclusions



- We present the Social Digital Twinner,
  - an advanced and democratized platform
  - designed to build on-demand social digital twins
  - for delivering deep insights into societal challenges at various scales,
  - enhancing decision-making and policy development.
- Twinner achieves greater realism, flexibility, and scalability thanks to the combination of cutting technologies like ABM, ML and LLMs.
- We implemented of PoC for studying the school drop-out in Kragerø.
- The techniques and methodologies employed in Twinner have not been formally evaluated yet.

# Thank you.

