

# DATATOPIA

## The World of Unlimited Industrial Data and Exploration



Michael Siegl<sup>1</sup>, Aysenur Gilik<sup>1</sup>, Vaishali Dhanoa<sup>1</sup>, Jan Holzweber<sup>1</sup>, Michael Haslgrübler<sup>1</sup>,  
Muaaz Abdul-Hadi<sup>2</sup>, Elisei Ember<sup>2</sup>, Tanveer Ali Ahmad<sup>2</sup>, Belgin Mutlu<sup>2</sup>, Christina Pichler<sup>2</sup>

Pro2Future GmbH<sup>1,2</sup>

<sup>1</sup> Science Park 4, Altenberger Strasse 69, 4040 Linz

<sup>2</sup> Inffeldgasse 25F, 8010 Graz



## MOTIVATION & GOALS

Currently **deep learning** in industrial use cases is **held back** by the **security** concerns linked to production line data. One such concern is that confidential **process specific knowledge** can be inferred from raw sensor data. Thus, such data is closely held. Furthermore, it is **seldom annotated**, which greatly restricts supervised learning methods.

A solution to data scarcity is using **generative methods** to enhance available datasets. While generating data solves data availability, the **contained information** will still be similar. In order to increase a datasets information content, we aim to include **domain expertise**. The generated data will be presented in an interactive framework. Using this framework, the employed generative methods can be enriched with **human experience**.

### Project FactBox

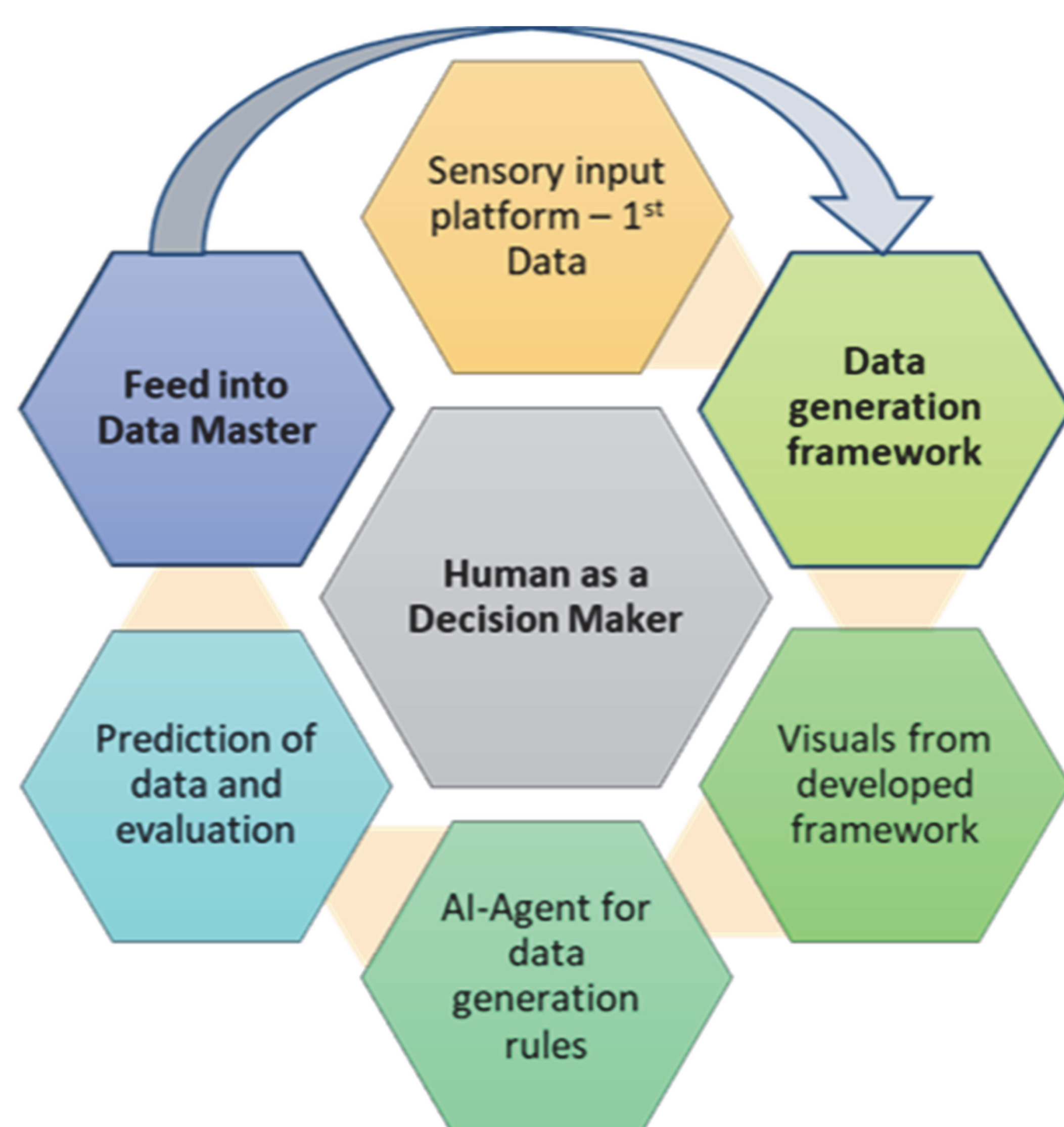
**Project Name** Datatopia  
**Project ID** StratP SX200  
**Duration** 10 Months

**Area 1**  
Perception and Aware Systems  
**Area 2**  
Cognitive Robotics & Shopfloors  
**Area 3**  
Cognitive Decision Making  
**Area 4.1**  
Cognitive Products  
**Area 4.2**  
Cognitive Production Systems

**Project Lead**  
Christina Pichler

## FRAMEWORK

Starting from **sensor input** data, generative methods are **learned**. In order to improve the contained information, rules for generation are set. Then generation can be **reiterated** upon to improve results further.



## APPROACH

### Generative:

- Autoregressive (GAN, VAE)
- Generative modelling (Gaussian copula)

### Interactive:

Including domain expertise in the quality control of generated data

### Secure:

- Masking process specifics
- Exposing minimal data
- Using black box models

## CONTRIBUTION

### Scientific contribution

- Generative AI (Time Series, Ultra-Wide Band)
- General representations for industry specific data

### Economic contribution

- Increase industrial data privacy and availability
- Establish process security

**Contact:** Michael Siegl, BSc, Pro2Future GmbH, michael.siegl@pro2future.at

**Acknowledgement:** This work was supported by Pro²Future (FFG, 881844).