problem

idea

dataset

experiments

conclusions

Unleashing the Autoconversion* Rates Forecasting: Evidential Regression from Satellite Data

Maria Carolina Novitasari (UCL)
Johannes Quaas (U. Leipzig)
Miguel R. D. Rodrigues (UCL)

*) Rain formation in liquid clouds.

in collaboration with

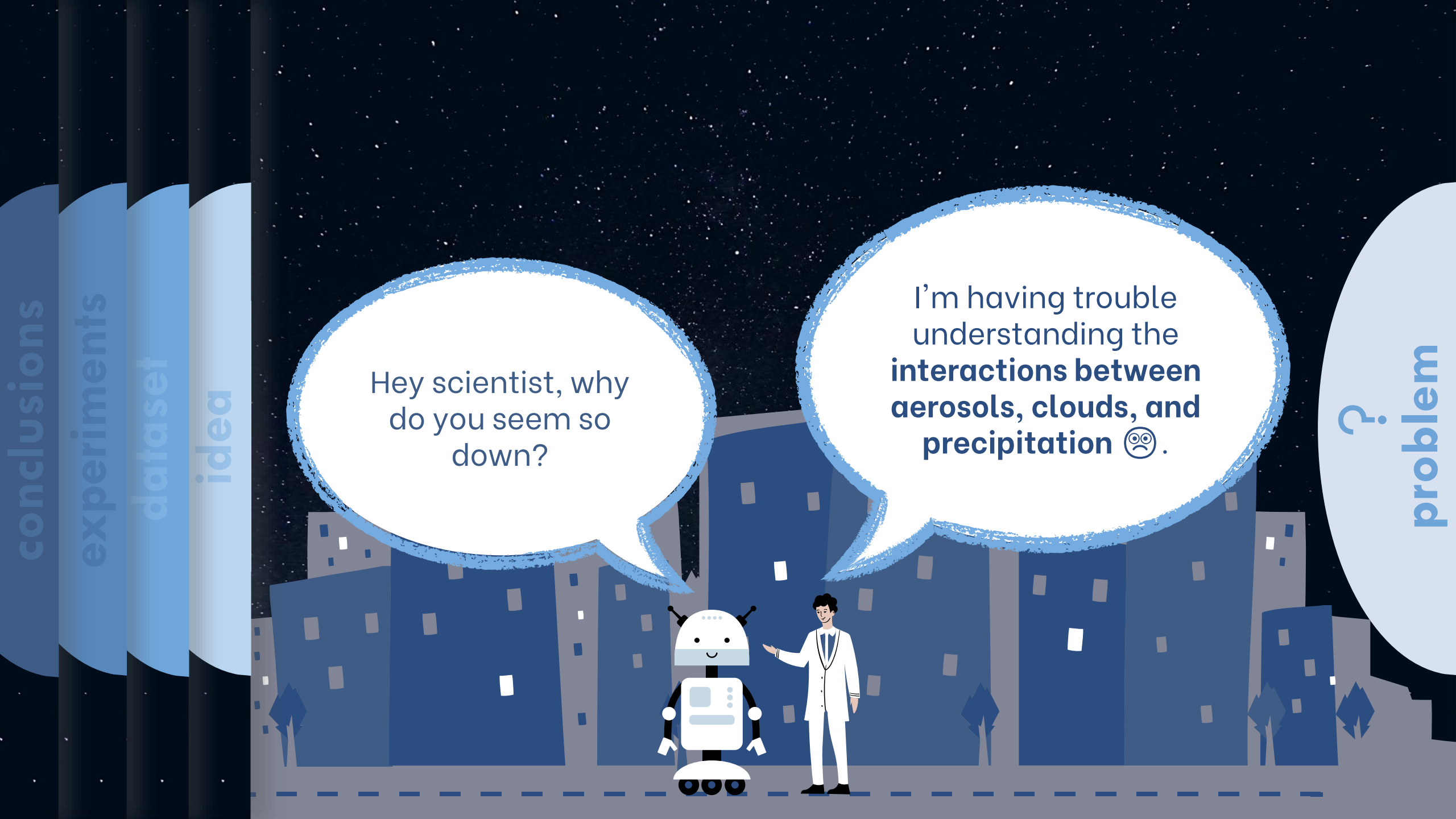
UNIVERSITÄT LEIPZIG

iMIRACLI

This research receives funding from the European Union's Horizon 2020 research and innovation programme under Marie Skłodowska-Curie grant agreement No 860100 (iMIRACLI).

Climate Change AI

NEURAL INFORMATION PROCESSING SYSTEMS


conclusions
experiments
dataset
idea
?
problem
Hey scientist, why do you seem so down?
I'm having trouble understanding the **interactions between aerosols, clouds, and precipitation** 😟.

conclusions

experiments

dataset

idea

?

problem

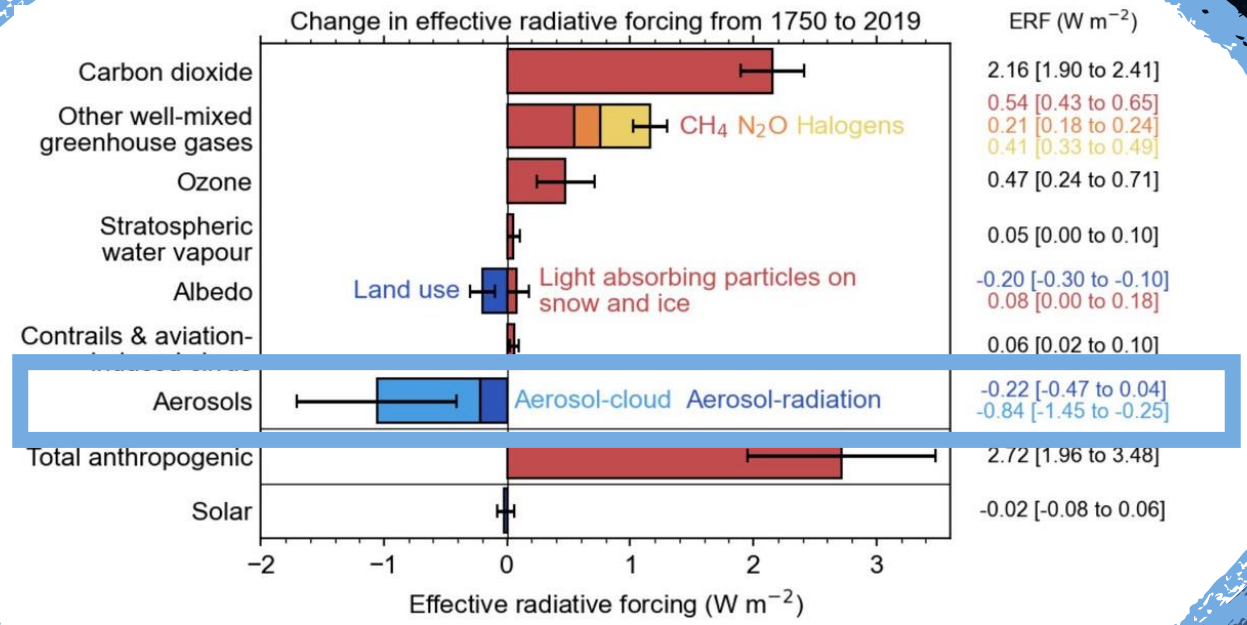
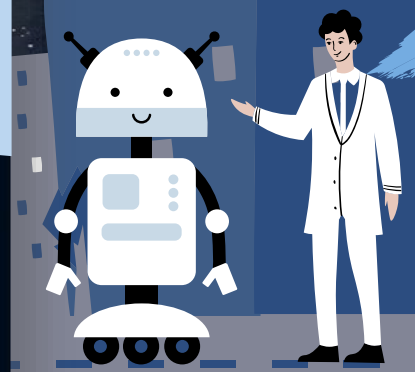
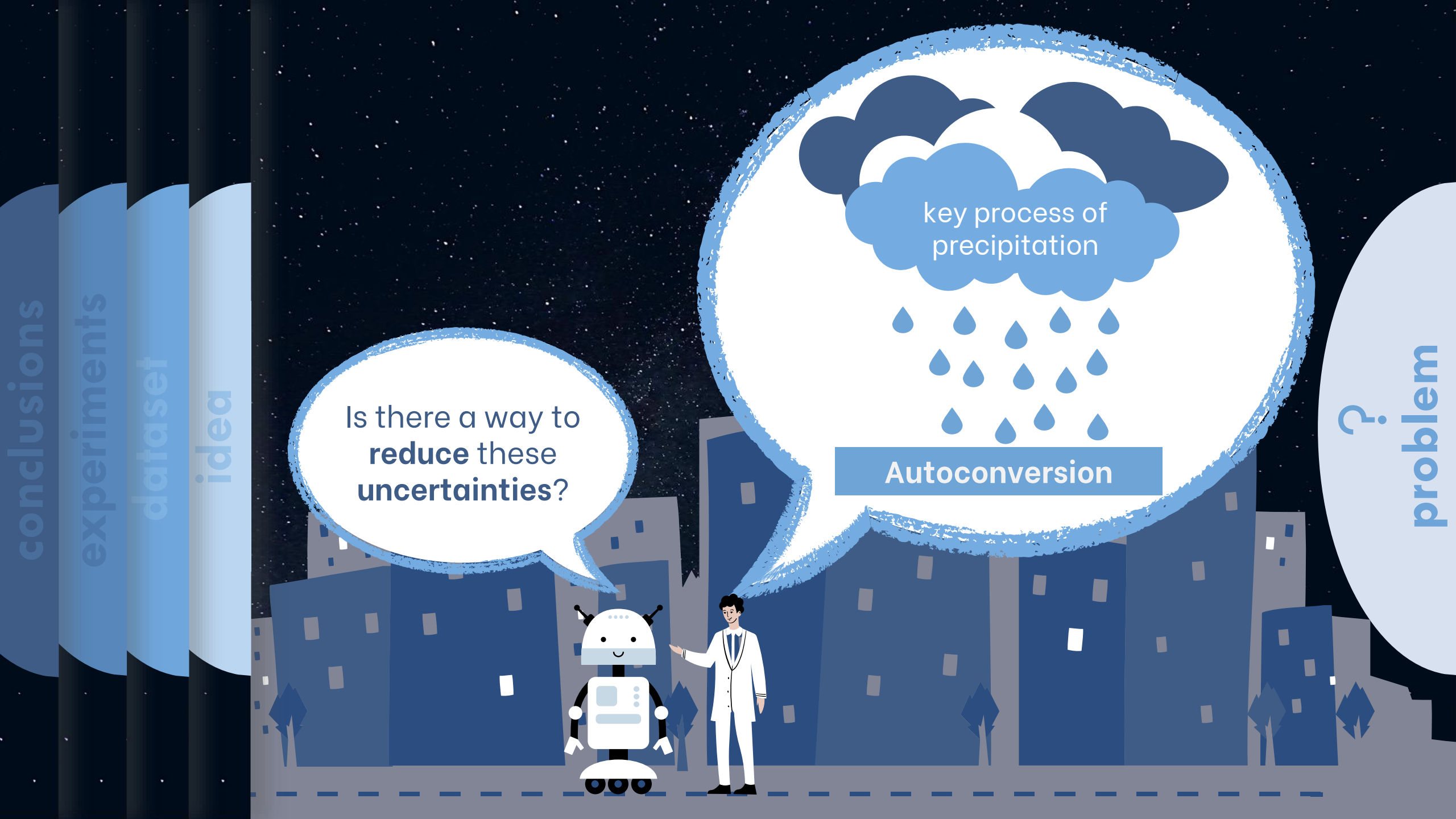


Image obtained from Figure 7.6 of Intergovernmental Panel on Climate Change (IPCC) 2021 AR6 Report of Working Group 1.




conclusions
experiments
dataset
idea
?
problem
key process of precipitation
Autoconversion
Is there a way to **reduce** these **uncertainties**?

conclusions

experiments

dataset

idea

?

problem

Don't you have **atmospheric simulation** that can help with this?

ICON-LEM

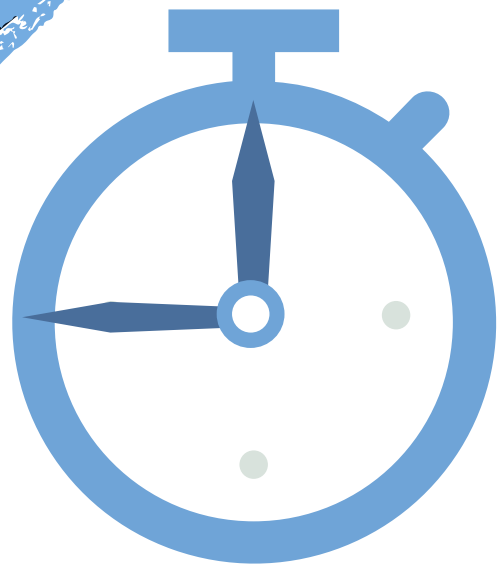

conclusions
experiments
dataset
idea
?
problem
So, what is the **issue**?

conclusions

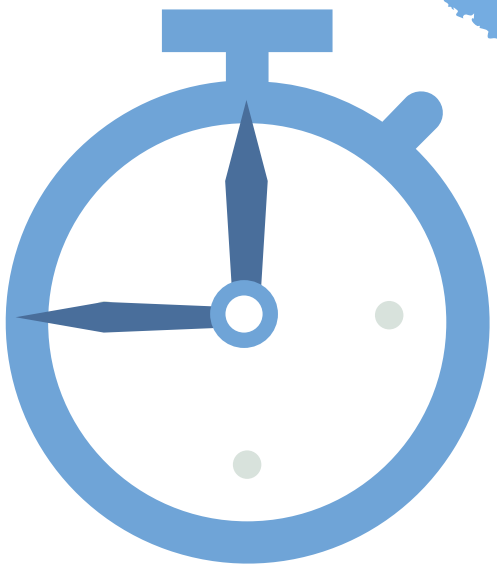
experiments

dataset

idea



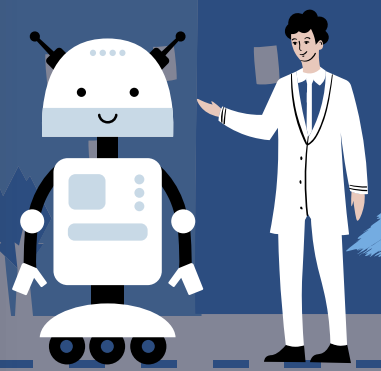
1 simulated hour

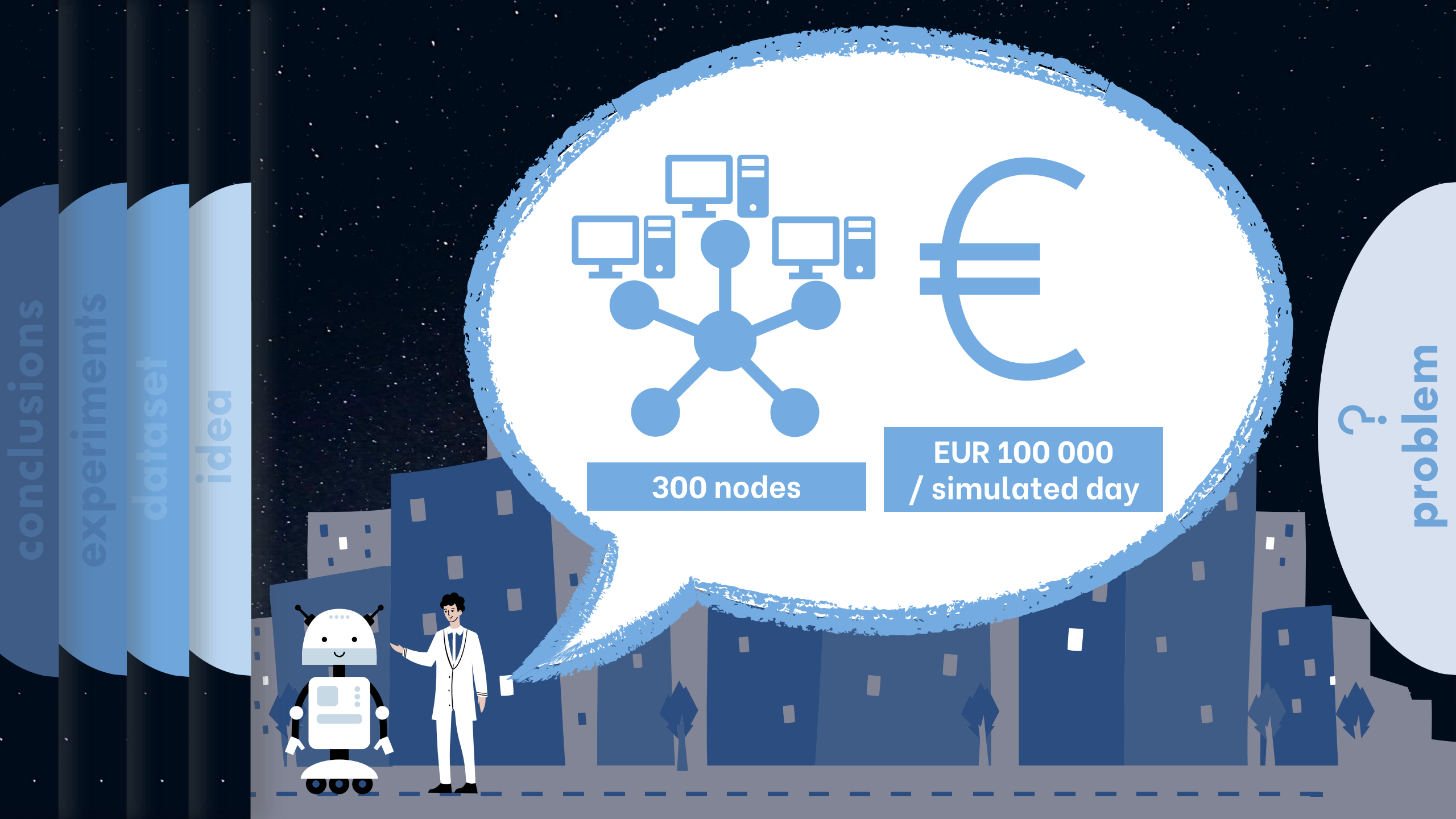


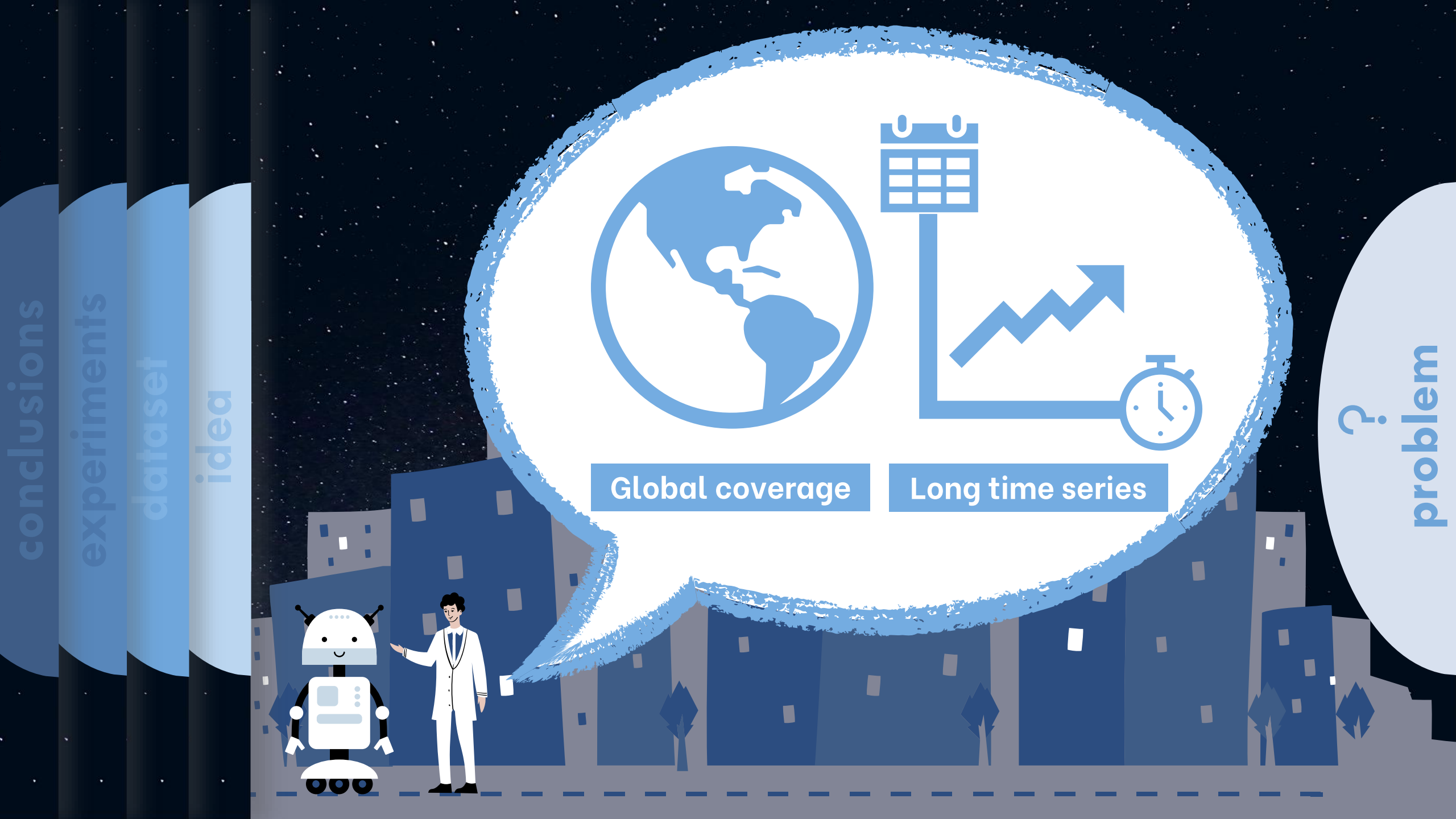
Real running time: 13 hours

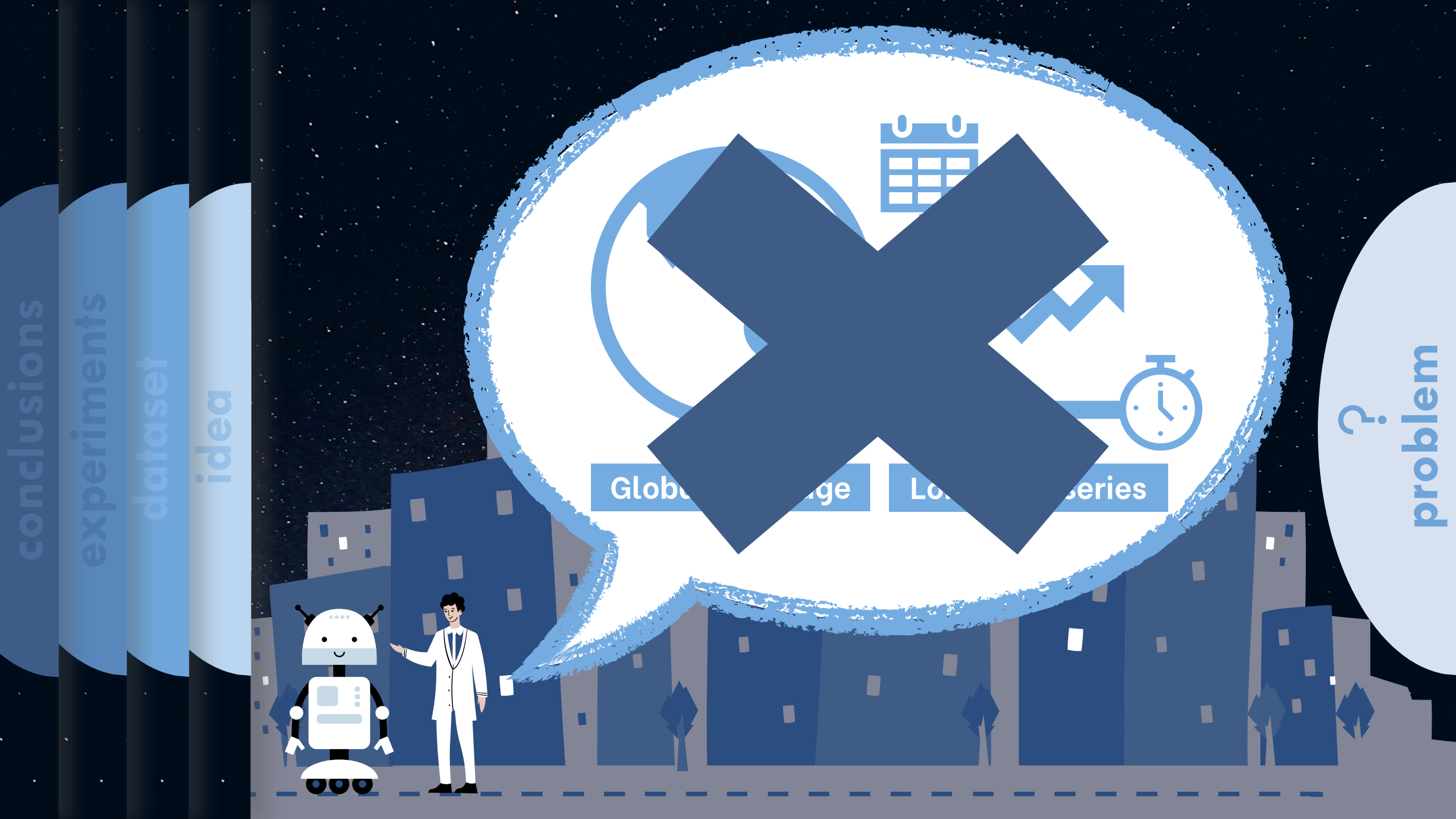
?

problem




conclusions
experiments
dataset
idea
?
problem
300 nodes
EUR 100 000
/ simulated day


conclusions
experiments
dataset
idea
problem
?
Global coverage
Long time series


conclusions
experiments
dataset
idea
?
problem
Globe ge
Lo eries

conclusions

experiments

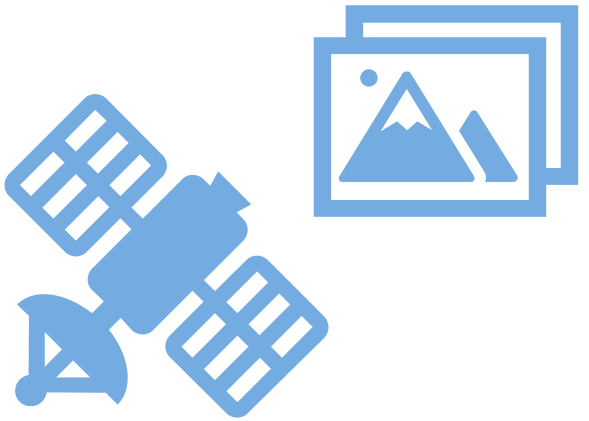
dataset

idea

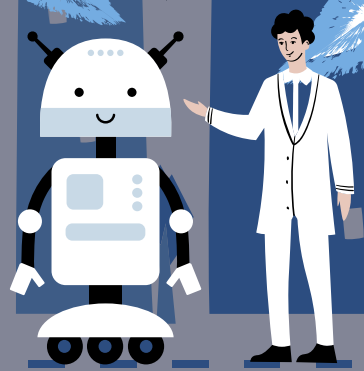
problem

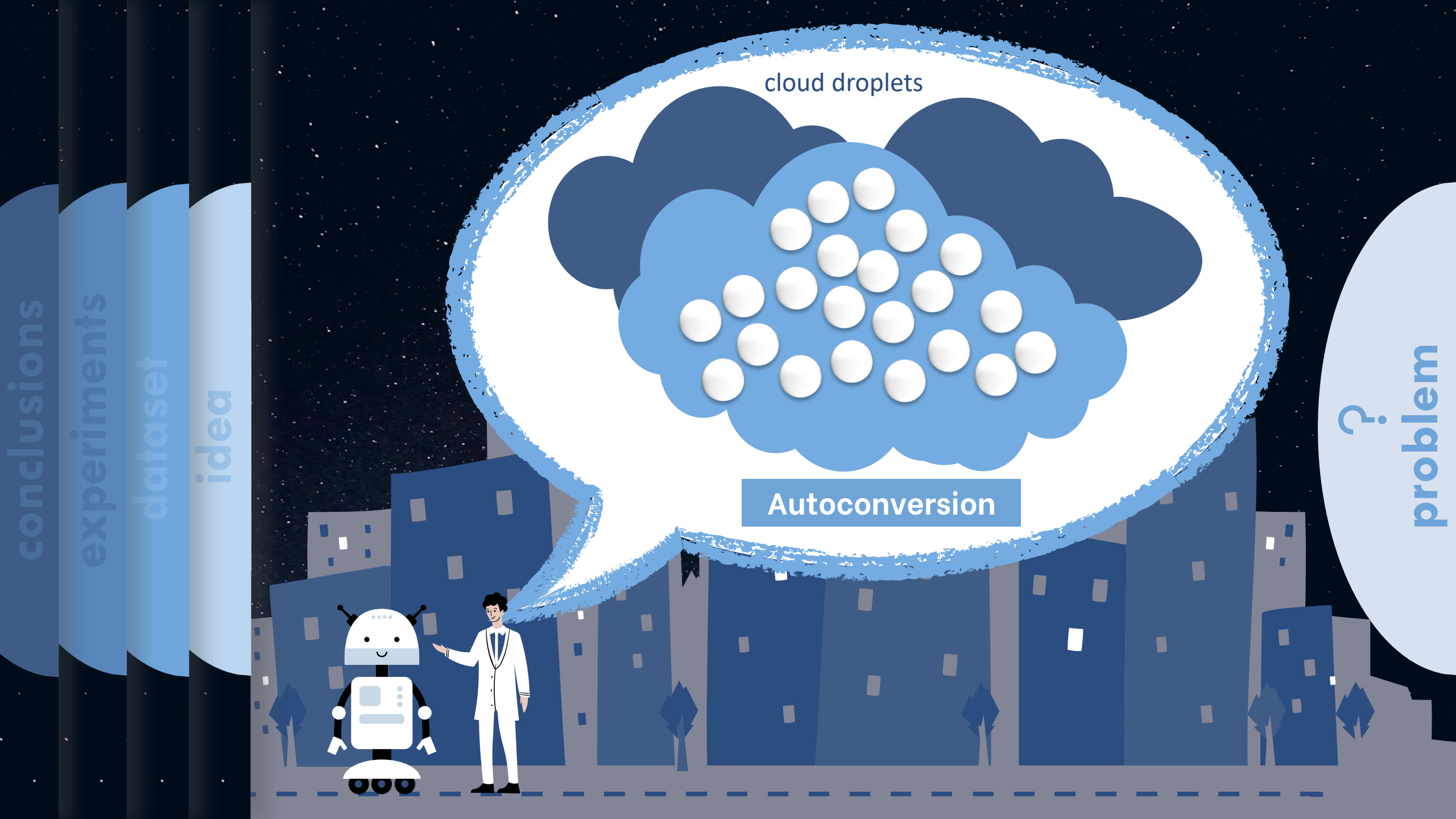
I'm a **machine learning** you know! Maybe I can help with my magic! 😉 Btw, what is the **autoconversion**?

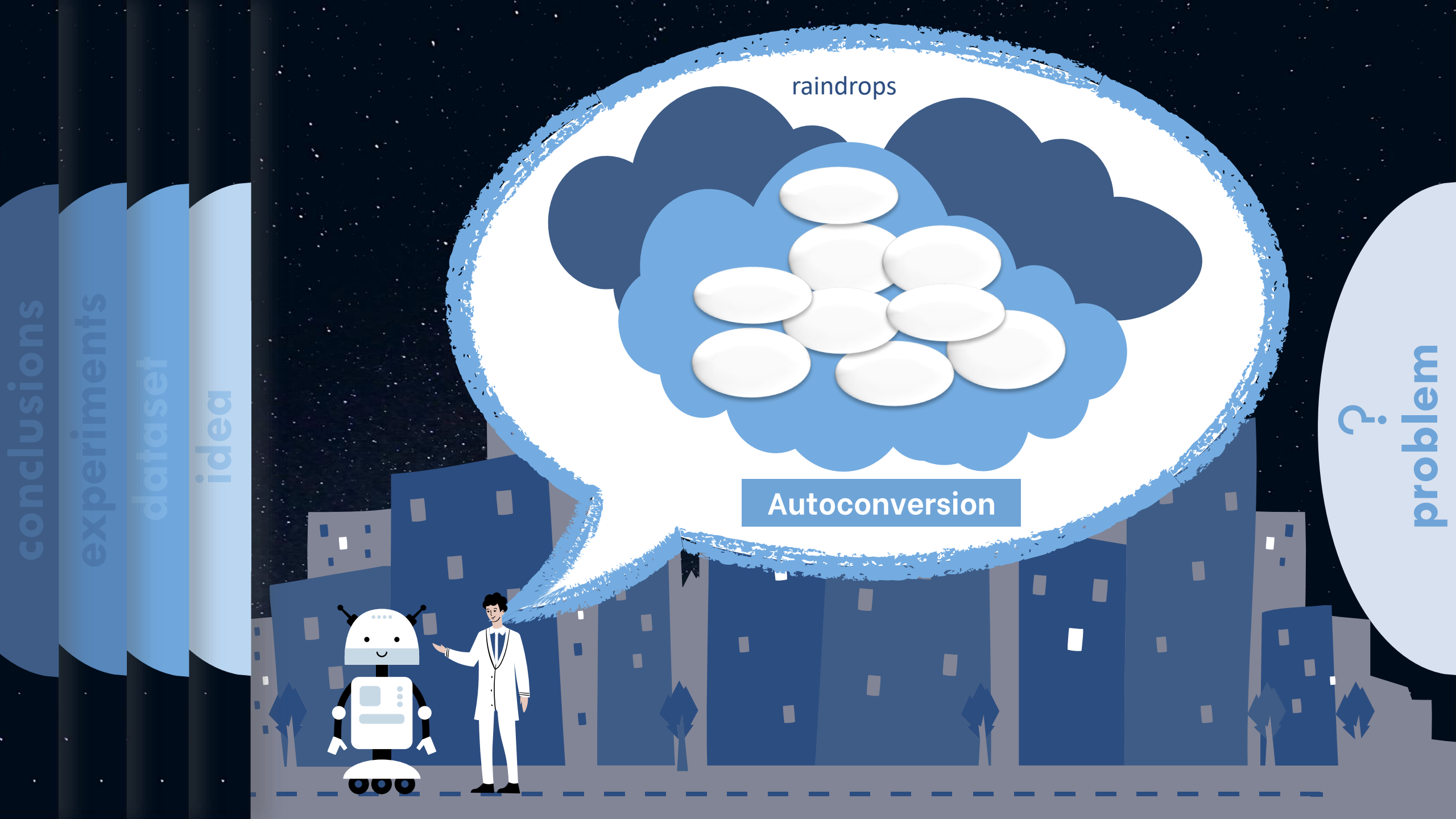
long-term, global coverage

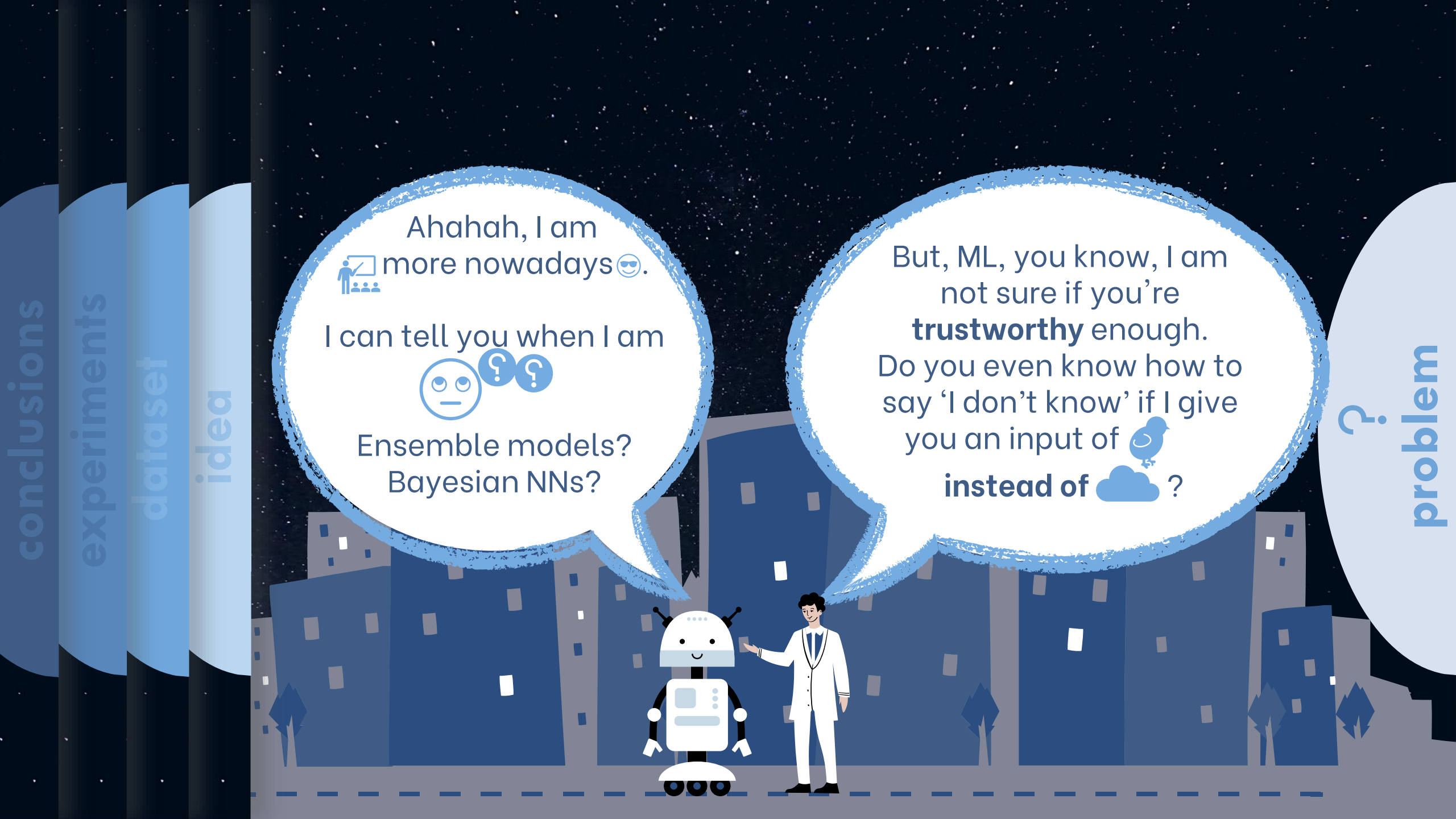


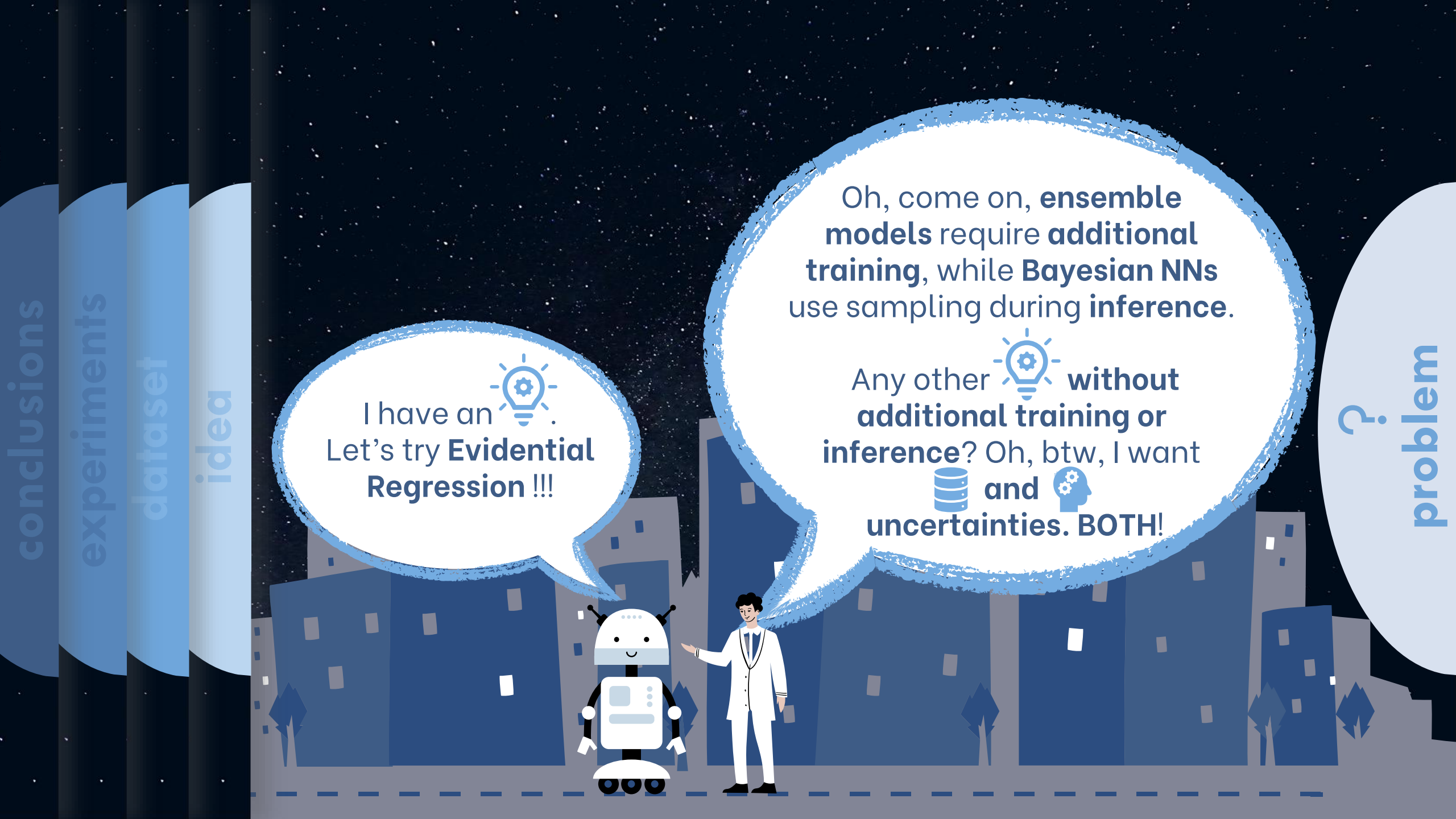
Satellite data




cloud droplets
Autoconversion
conclusions
experiments
dataset
idea
?
problem


raindrops
Autoconversion
conclusions
experiments
dataset
idea
?
problem


conclusions
experiments
dataset
idea
Ahahah, I am more nowadays 😎.
I can tell you when I am
Ensemble models?
Bayesian NNs?
But, ML, you know, I am not sure if you're **trustworthy** enough.
Do you even know how to say 'I don't know' if I give you an input of
instead of ?
problem


conclusions
experiments
dataset
idea
problem
?
I have an .
Let's try **Evidential Regression** !!!
Oh, come on, **ensemble models** require **additional training**, while **Bayesian NNs** use sampling during **inference**.
Any other **without additional training or inference**? Oh, btw, I want and **uncertainties. BOTH**!

problem

idea

NN model

Train data on NN

Input:
cloud effective radius

Output:
autoconversion rates + evidential parameters $(\gamma, v, \alpha, \beta)$

Atmospheric Simulation

dataset

experiments

conclusions

problem

idea

dataset

experiments

conclusions

Satellite Data

NN model

autoconversion rates + evidential parameters $(\gamma, v, \alpha, \beta)$

uncertainty estimation

Evidential Regression

Total loss = maximizing evidence + regularizing evidence

$$\mathcal{L}_i(\boldsymbol{w}) = \mathcal{L}_i^{\mathrm{NLL}}(\boldsymbol{w}) + \lambda \cdot \mathcal{L}_i^{\mathrm{R}}(\boldsymbol{w})$$

Evidence Regularizer (Amini, 2020)

$$\mathcal{L}_i^{\mathrm{R}}(\boldsymbol{w}) = |y_i - \mathbb{E}\left[\mu_i\right]| \cdot \Phi = |y_i - \gamma| \cdot (2v + \alpha)$$

Evidence Regularizer (Meinert, 2023)

$$\mathcal{L}_i^{\mathrm{R}}(\boldsymbol{w}) = \left|\frac{y_i - \mathbb{E}[\mu_i]}{w_{St}}\right|^p \cdot \Phi = \left|\frac{y_i - \gamma}{w_{St}}\right|^p \cdot (v + 2\alpha)$$

Data Unc (Meinert, 2023)

$$u'_{\mathrm{al}} \equiv w_{\mathrm{St}} = \sqrt{\frac{\beta_i(1+\nu_i)}{\alpha_i \nu_i}}$$

Model Unc (Meinert, 2023)

$$u'_{\mathrm{ep}} \equiv \frac{u_{\mathrm{ep}}}{u_{\mathrm{al}}} = \frac{1}{\sqrt{\nu_i}}$$

problem

idea

dataset

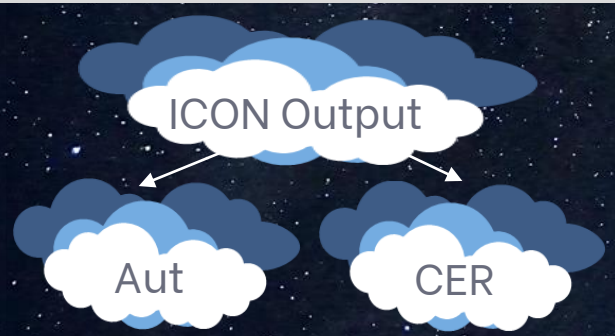
experiments

conclusions

Atmospheric Simulation Model (ICON)

LEM Germany – 02 May 2013
156 m regridded to 1 km
9:55am to 1:20pm local time

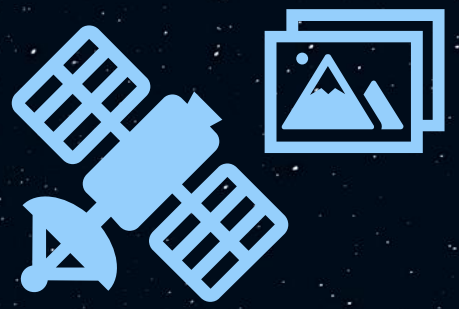
NWP Holuhraun – 2 Sept 2014
2.5 km horizontal resolution
11am local time



Satellite Simulator (Cloud-top ICON)

Match with MODIS Cloud Product Level 2

Satellite Observation (MODIS)



MODIS Cloud Product Level 2

CER: Cloud Effective Radius
Aut: Autoconversion rates
CT: Cloud-top

Autoconversion on Simulation Model (ICON)

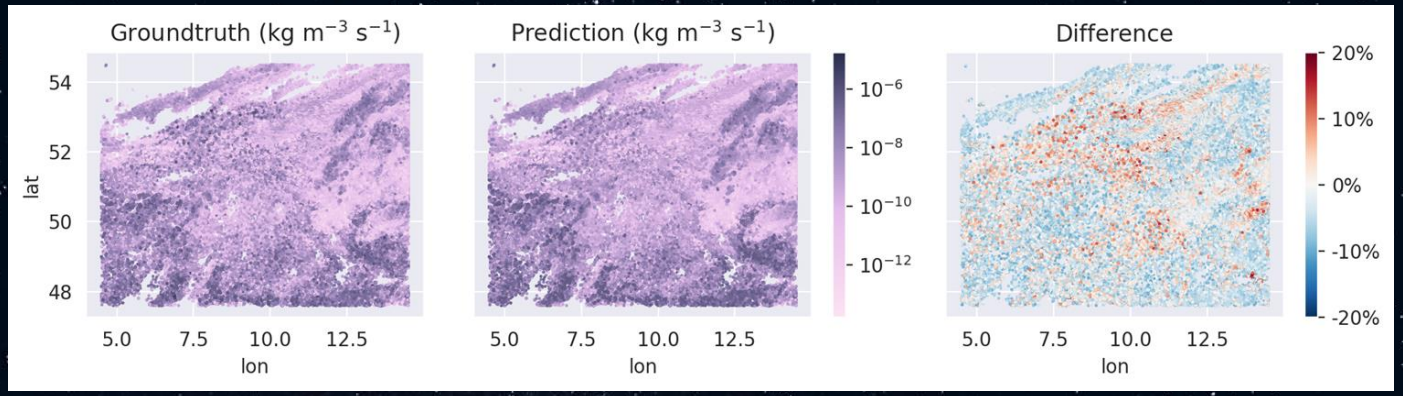
Relatively good outcomes for all scenarios

Dataset / Testing Scenario	R^2	MAPE	RMSPE	SSIM	PSNR
ICON-LEM Germany (1 km)	90.54%	9.14%	11.27%	90.11%	26.30
Cloud-top ICON-LEM Germany (1 km)	89.88%	10.86%	13.89%	89.96%	25.89
Cloud-top ICON-LEM Holuhraun (2.5 km)	84.39%	8.56%	10.73%	91.50%	25.81

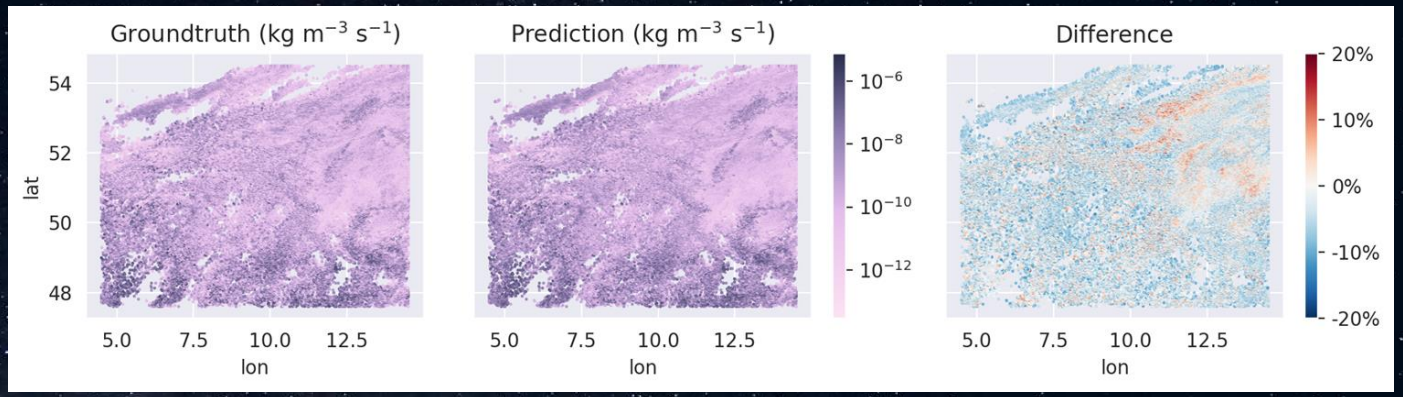
Autoconversion on Simulation Model (ICON)

Visual Representation of Scenario 1, 2, and 3

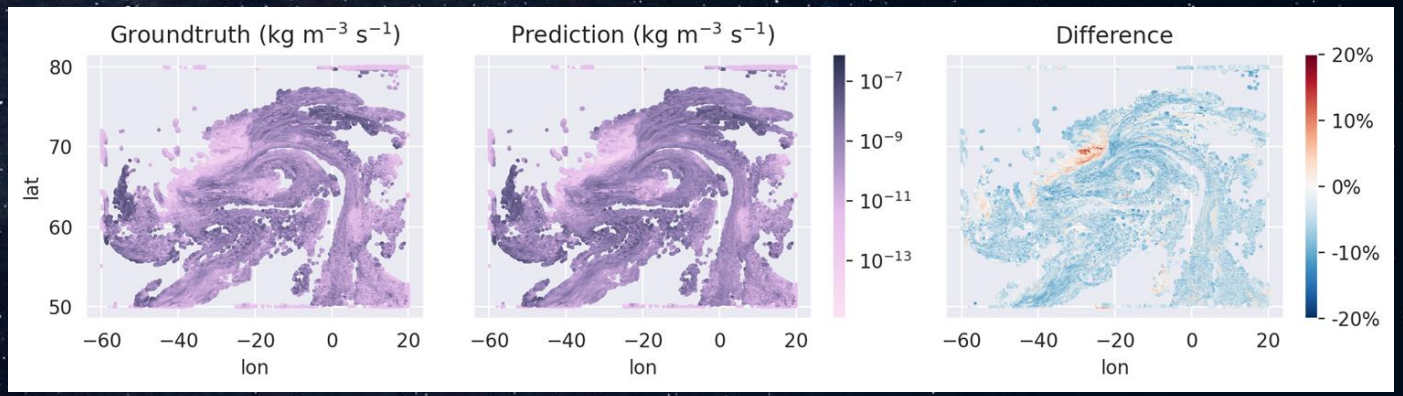
ICON-LEM Germany (1 km)



Cloud-top ICON-LEM Germany (1 km)



Cloud-top ICON-NWP Holuhraun (2.5 km)



conclusions

experiments

dataset

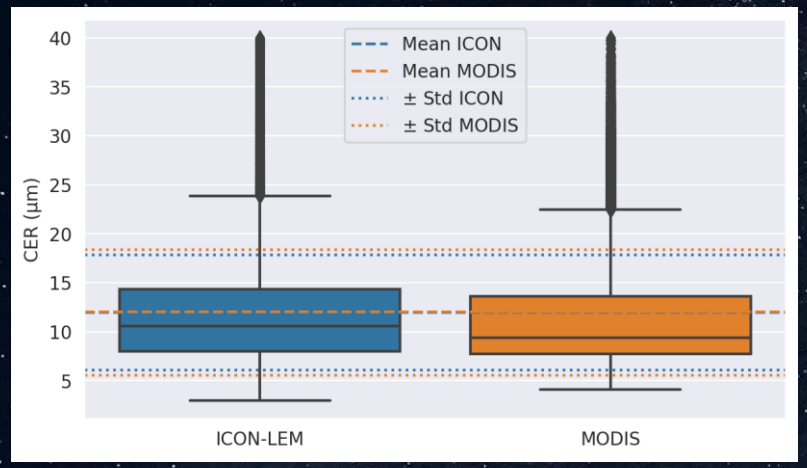
idea

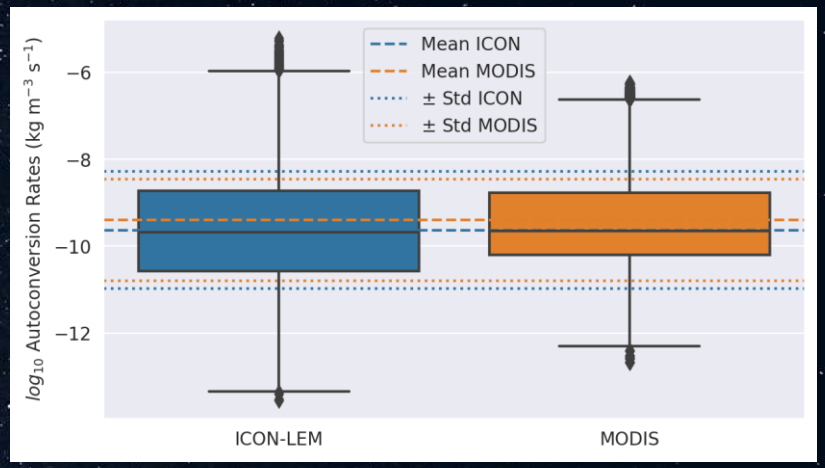
problem

Autoconversion on Satellite Observation (MODIS)

Statistical Concordance

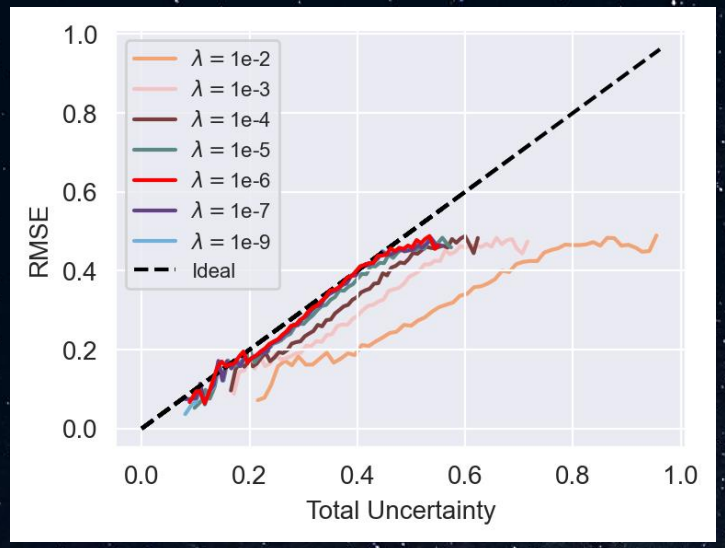
MODIS Aqua over Germany,
13:20 UTC
AoI = [5.87, 47.50, 10.00, 54.50]



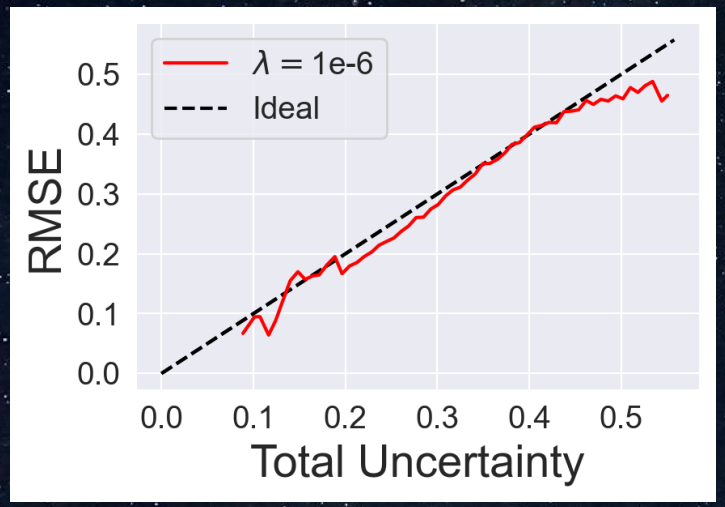


Mean, standard deviation, median, 25th and 75th percentiles of Cloud-top ICON-LEM Germany and MODIS: autoconversion rates (Aut)

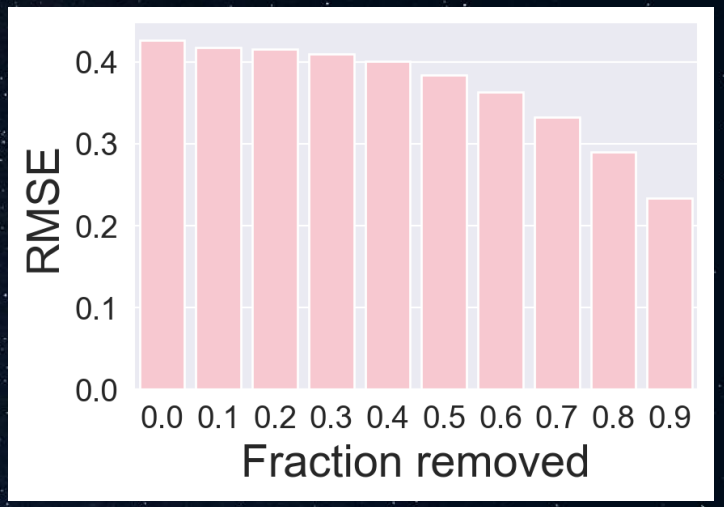
Evaluation of Uncertainty Estimation



Well-calibrated



Spread-skill plot



Discard test

conclusions

experiments

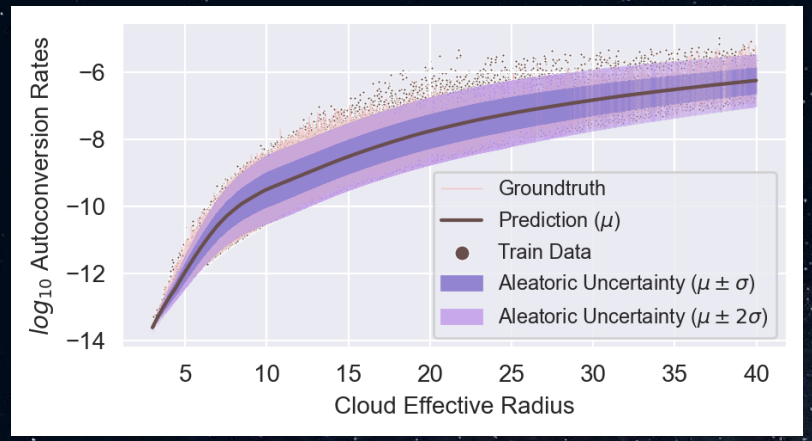
dataset

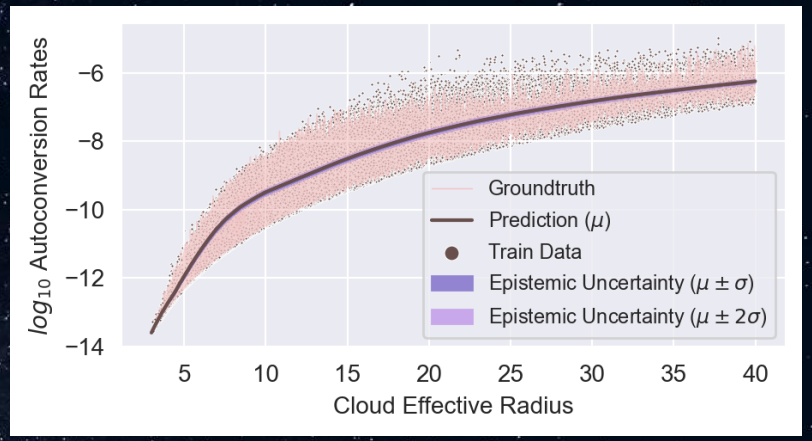
idea

problem

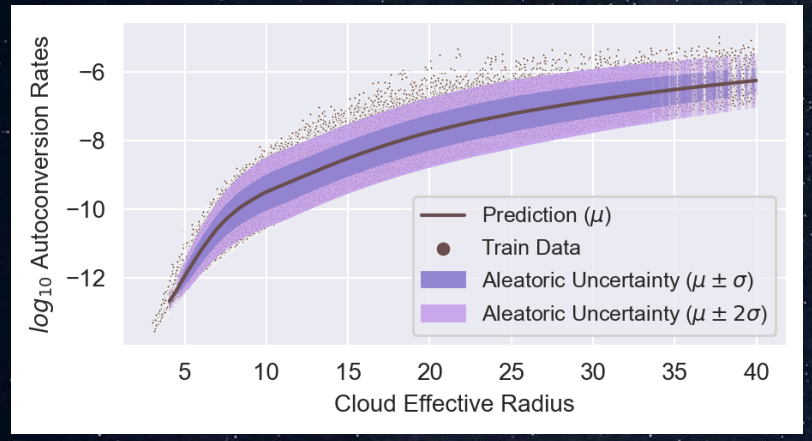
Uncertainty Estimation

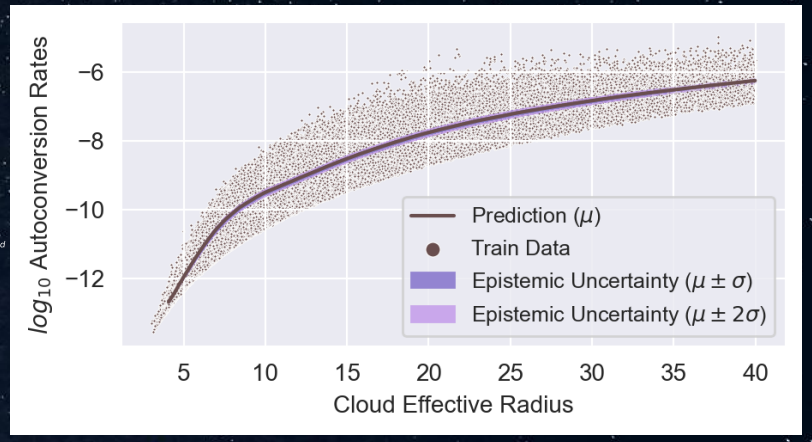
Simulation Model (ICON-LEM Germany) – 02 May 2013



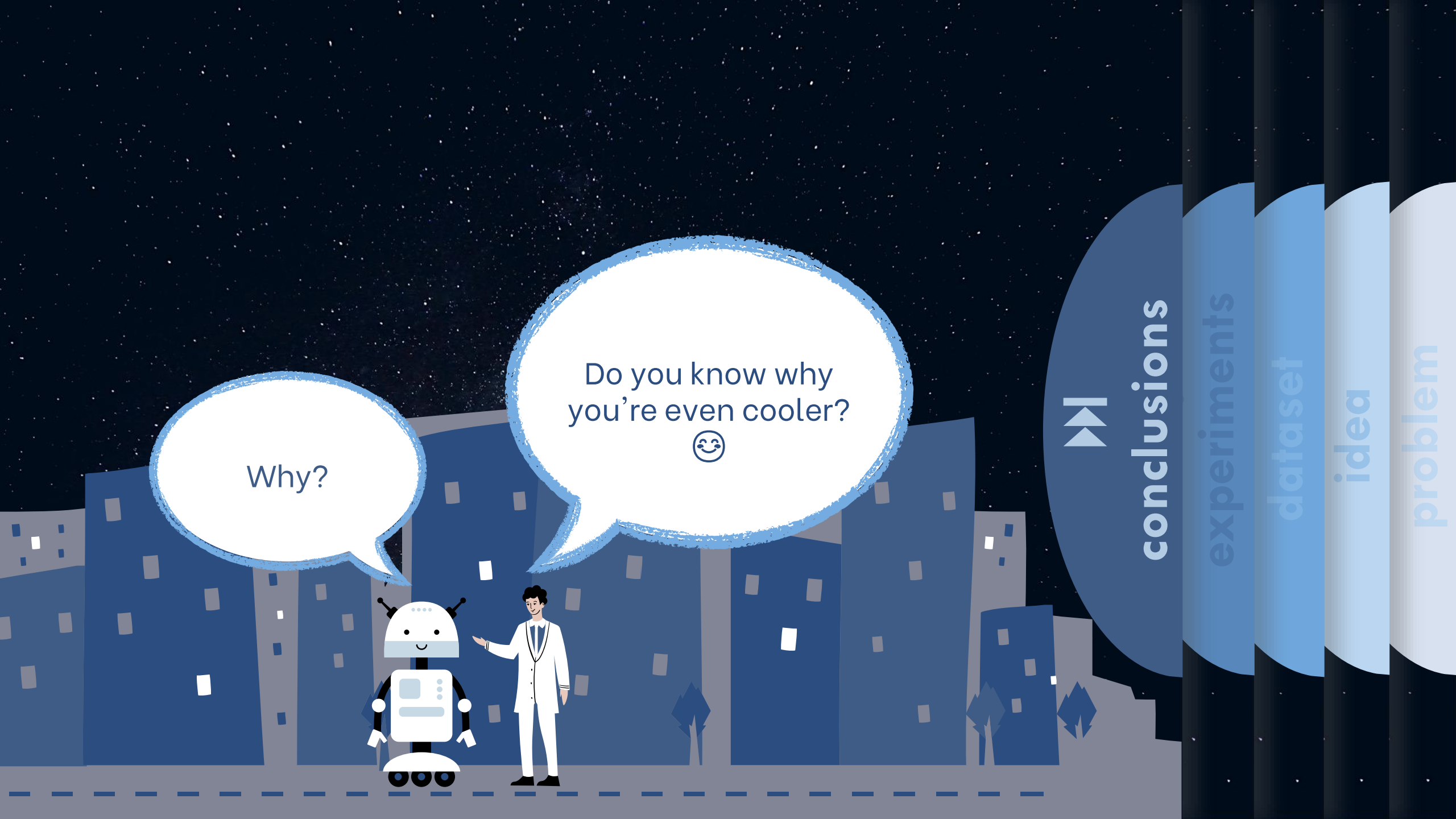


Satellite Data (MODIS Germany) – 02 May 2013

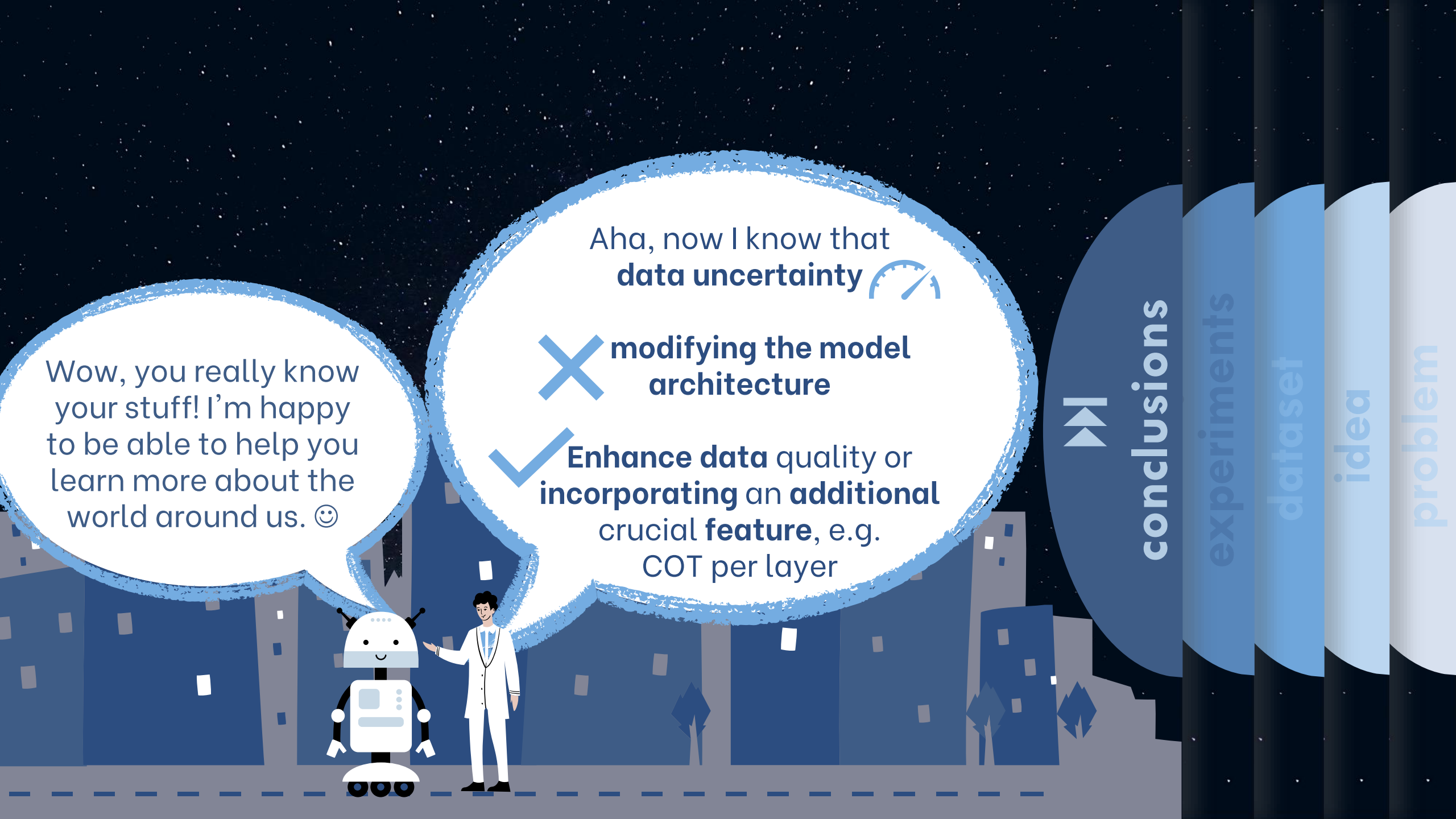





Hey ML, thanks so much for helping me **uncover the key process of precipitation directly from satellite data**. You're so cool!
You're very welcome! ☺
conclusions
experiments
dataset
idea
problem


Why?
Do you know why you're even cooler?
conclusions
experiments
dataset
idea
problem


I can't take all the credit though 😉.
Because you are **able to estimate** and **uncertainties.**
additional training/inference.
conclusions
experiments
dataset
idea
problem


Wow, you really know your stuff! I'm happy to be able to help you learn more about the world around us. ☺
Aha, now I know that **data uncertainty**
modifying the model architecture
Enhance data quality or **incorporating** an **additional** crucial **feature**, e.g. COT per layer
conclusions
experiments
dataset
idea
problem

Thank you! ☺

This research receives funding from the European Union's Horizon 2020 research and innovation programme under Marie Skłodowska-Curie grant agreement No 860100 (iMIRACLI).