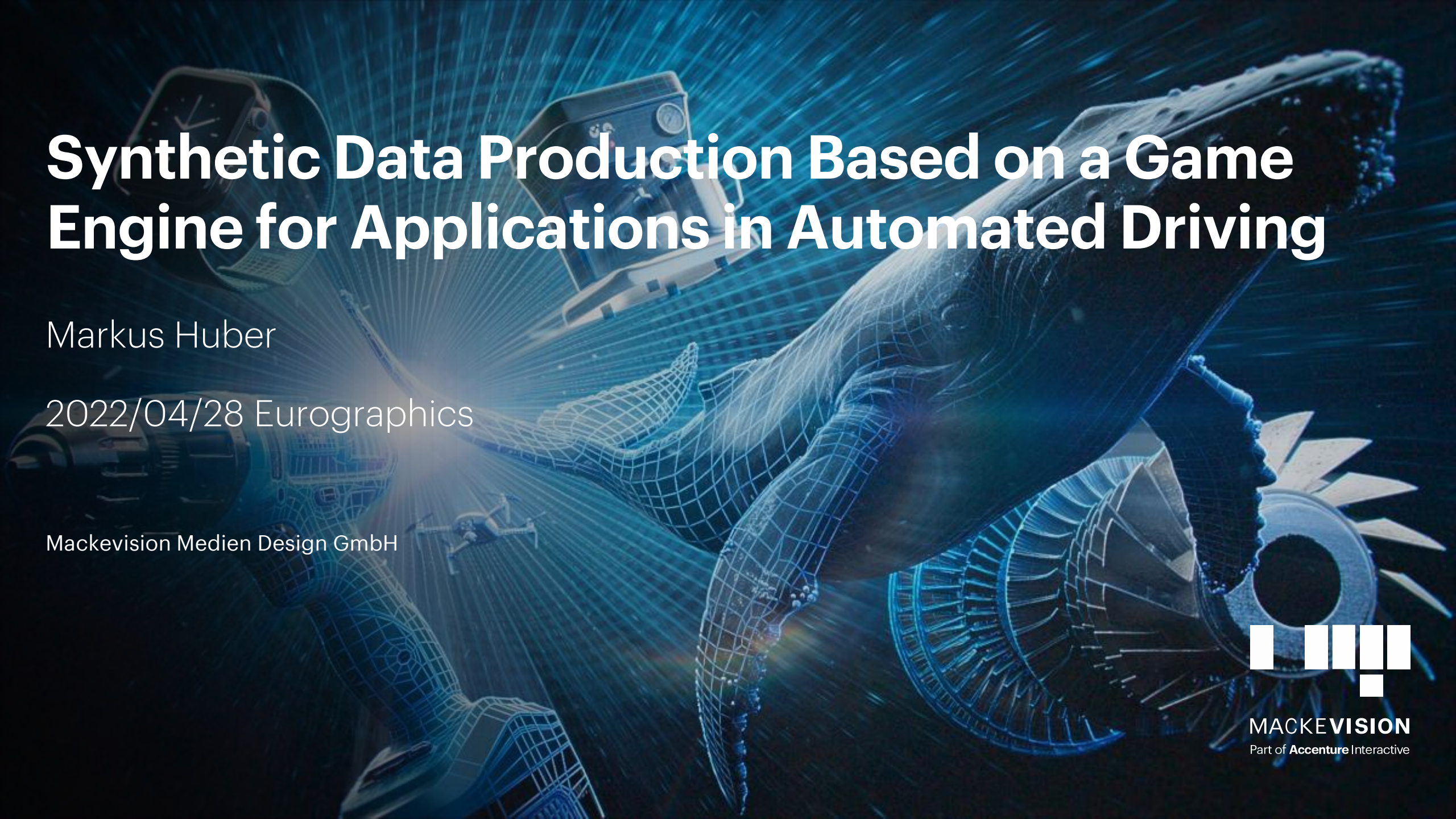
Synthetic Data Production Based on a Game Engine for Applications in Automated Driving

Markus Huber

2022/04/28 Eurographics

Mackevision Medien Design GmbH



MACKEVISION
Part of Accenture Interactive

Synthetic Data for Automated Driving

Synthetic image data for AI/ML applications with digital twin approach

- “Collect” data in virtual worlds

Benefits

- Systematic (reproducible) data generation
- Automated generation of ground truth and meta data
- Focus on rare data or situations that are difficult to capture
- Scalability
- Easier handling of data privacy

Synthetic Data for Automated Driving

Data Requirements

- Realistic sensor data
- Broad coverage of traffic situations
- Compatibility to real-world data sets
- Parametrization and high degree of automation
- High variance
- Large amount of detailed content
- …

Why Game Engines?

UNREAL ENGINE

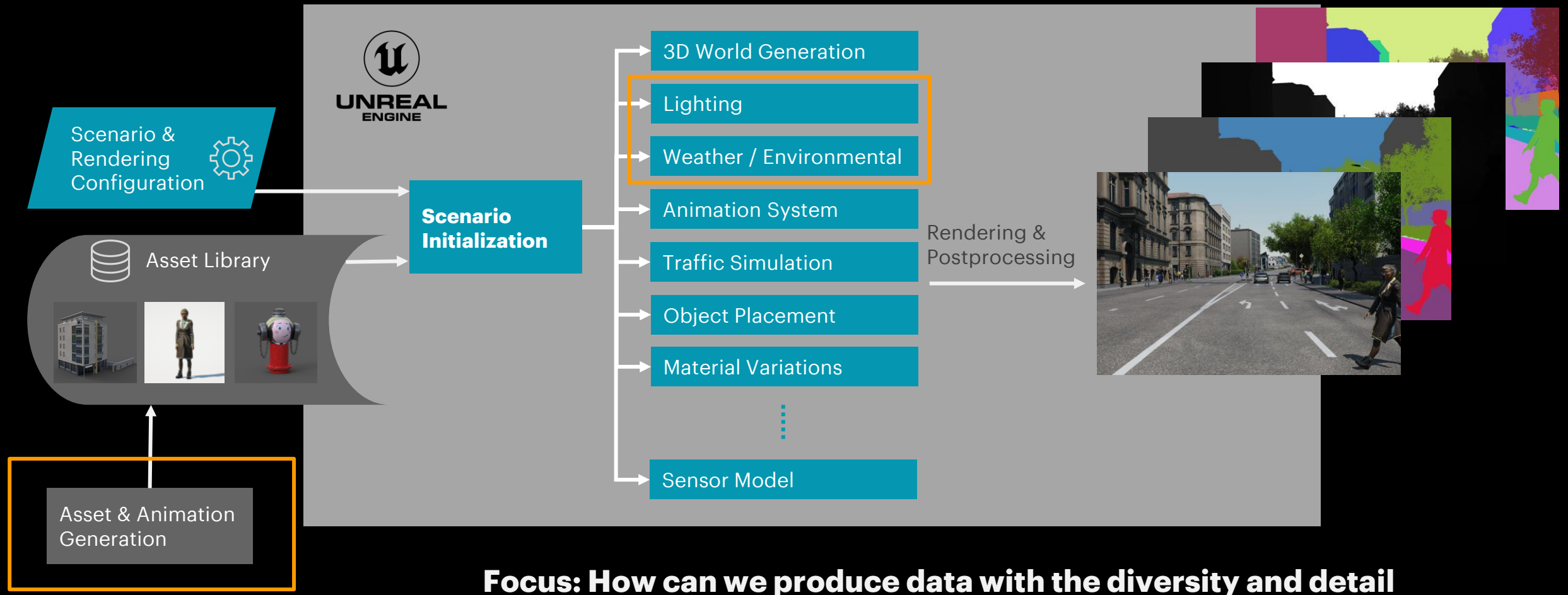
- Real-time engines: optimized for fast data production
- Proven capability to create complex and realistic worlds
- Features for dynamic content and objects
- Powerful material system
- High quality lighting system (ray tracing)
- Multiple flexible APIs (UE4: Blueprints, C++, Python)

- Development & application within joint research project “KI Absicherung”

KI ABSICHERUNG
Safe AI for Automated Driving

Supported by:
Federal Ministry for Economic Affairs and Climate Action
on the basis of a decision by the German Bundestag

Synthetic Data Generation



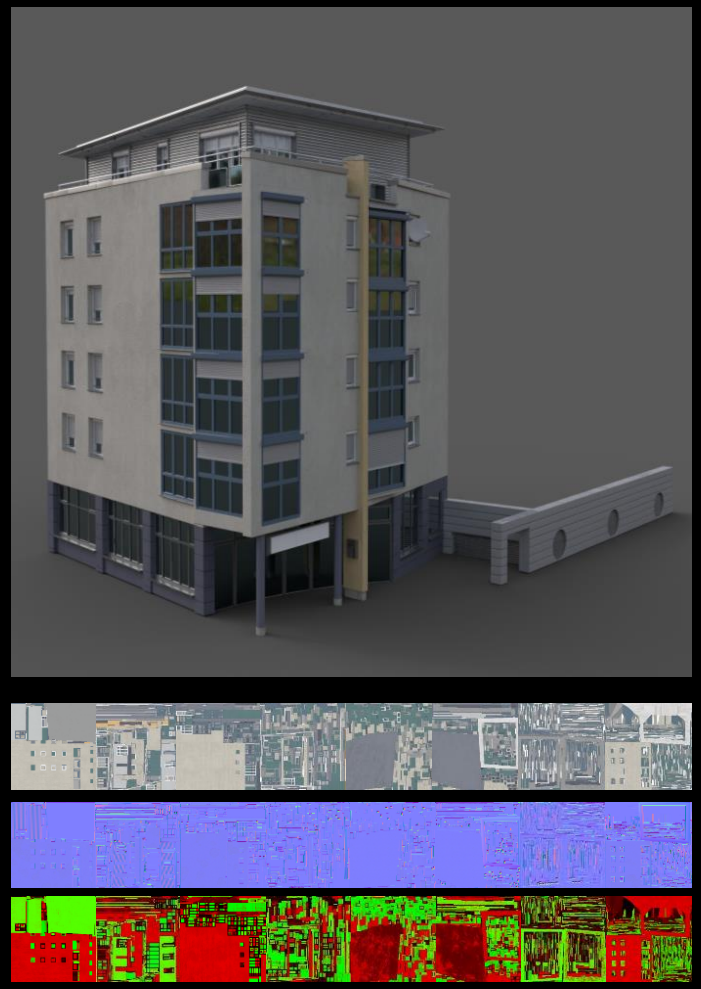
Focus: How can we produce data with the diversity and detail observed in real-world ?

Real-world Example

[Cityscapes, Cordt. et al. 2016]

Asset Generation & Animations

Asset Workflow



Game Assets with PBR Workflow

- Individual high-resolution textures with UDIMs
  → no repetitions, no tiling artifacts
  → sharp renderings
- Accurate material behavior under different lighting and weather conditions

- Compatibility to most render engines
- Can be used with
  - Artist created assets and materials
  - Photogrammetry workflow

Motion Capture Animations for Human Models

- Realistic motion for typical pedestrian movements
- Dedicated scenarios
- Typical as well as untypical and rare poses
- Interaction between persons
- Interaction with objects

Motion Capture Animations for Human Models

Motion Capture Animations for Human Models

Lighting & Weather

Procedural Sky

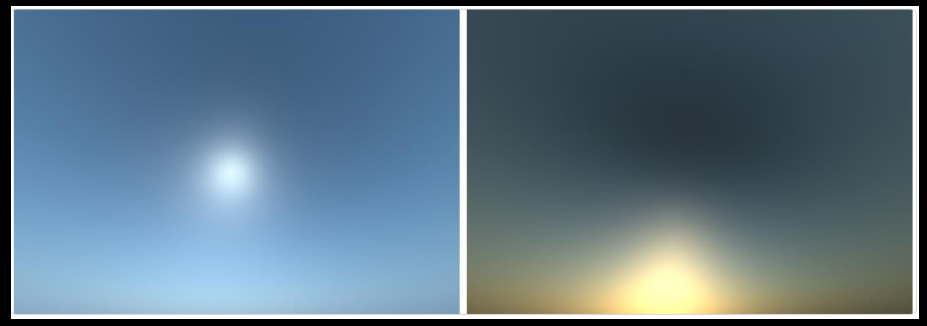
HDRI for skies

- Values in HDRIs unclear (possibly incorrect) if not calibrated
  → adjust brightness (guessing or by reference)
- If sun present in HDRI could have artifacts from camera/lens
  → (possibly error-prone) painting needed

Physical sky model

- Known (correct) values which is a good foundation for global illumination (GI) and reflections
- No artifacts around sun
- Without clouds can look boring

HDRI from http://noemotionhdrs.net/



Procedural sun

Procedural Clouds

Volumetric clouds

- Physical-based light and reflections with exciting scenery
- Artistic “look” control
- Arbitrary number of lighting situations controlled by random seed
- Easy and fast art direction when used in a real-time engine compared to using HDRIs
- Animated clouds possible

Materials: Wetness

Results, Next Steps, & Summary

Visual Results

Camera settings and light intensities from reference shoot

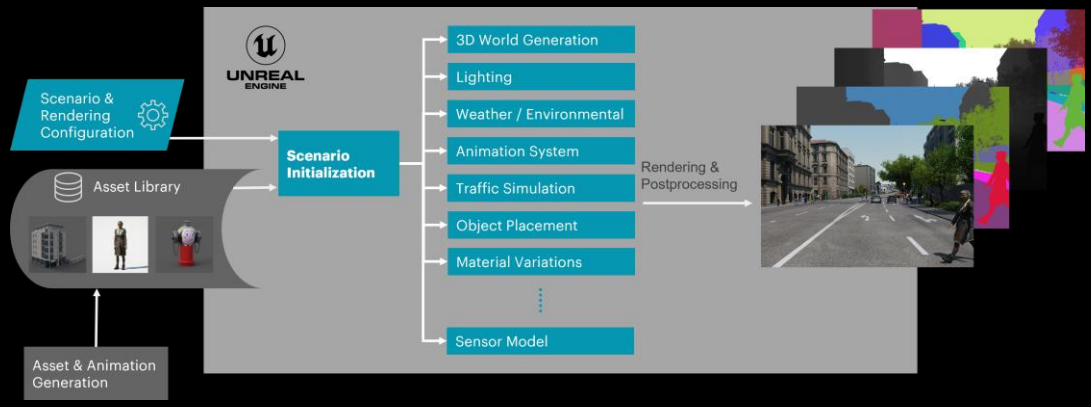
Varying Environments

Realtime weather system, volumetric clouds, realtime raytracing

What’s missing / next

Other Data Production Modules

- Procedural environments
- Sensor models and effects
- Error/defect generation
- Material measurements



Analysis, Evaluation, and ML Applications

- Measure realism and fidelity, comparison to real-world data
- Analysis from ML perspective
- Is there a “reality gap” and how to deal with it?
  - Adapt data production
  - Domain adaption
  - Style transfer

[Cityscapes, Cordt. et al. 2016]

Summary

Synthetic Image Data with Game Engines for Automated Driving Applications

- Game Engines provide many features that fit perfectly for the requirements of this application
- During data preparation and data production, several components should be exploited towards the goal of **diversity and detail** observed in real-world
  - Large selection of highly detailed assets
  - Advanced lighting
  - Environmental effects

Synthetic Data for Automated Driving

Data Requirements

- Realistic sensor data
- Broad coverage of traffic situations
- Compatibility to real-world data sets
- Parametrization and high degree of automation
- High variance
- Large amount of detailed content
- …

Why Game Engines?

UNREAL ENGINE

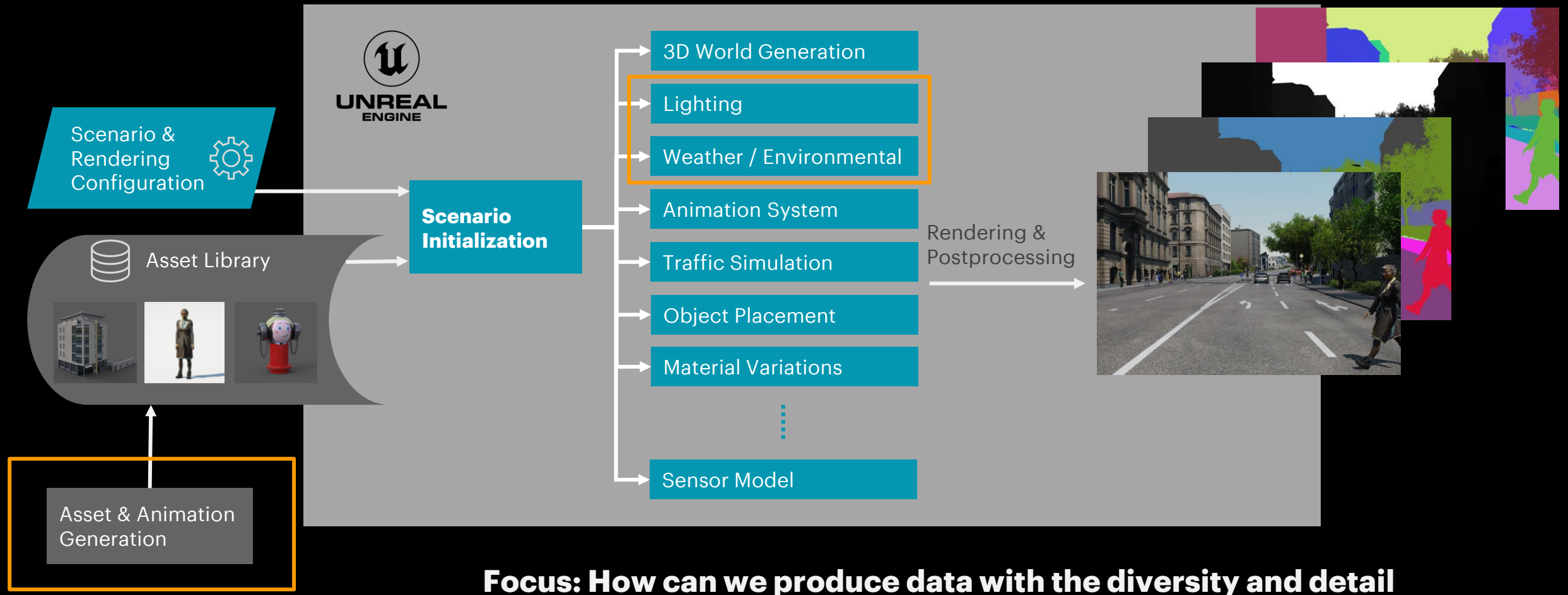
- Real-time engines: optimized for fast data production
- Proven capability to create complex and realistic worlds
- Features for dynamic content and objects
- Powerful material system
- High quality lighting system (ray tracing)
- Multiple flexible APIs (UE4: Blueprints, C++, Python)

- Development & application within joint research project “KI Absicherung”

KI ABSICHERUNG
Safe AI for Automated Driving

Supported by:
Federal Ministry for Economic Affairs and Climate Action
on the basis of a decision by the German Bundestag

Synthetic Data Generation



Focus: How can we produce data with the diversity and detail observed in real-world ?

Thank you!

Markus Huber

markus.huber@mackevision.com