AUTODESK UNIVERSITY

Using generative design and machine learning for faster analysis feedback

Varvara Toulkeridou Autodesk Sr. Research Engineer

Spyridon Ampanavos Harvard University Doctoral Candidate

Michael Floyd Autodesk AEC Sustainability Strategy Manager

Vishal Vaidhyanathan Autodesk Computational BPA Researcher

About the speakers

Michael Floyd

AEC Sustainability Strategy Manager at Autodesk. Environmental technologist working to bring better sustainability solutions to life through Autodesk's products. 14 years experience in design & sustainability.

Varvara Toulkeridou

Sr. Research Engineer for Autodesk's AEC Generative Design group. Doctoral candidate in Computational Design at Carnegie Mellon University, researching how machine learning can augment parametric modeling tools for supporting design exploration.

Vishal Vaidhyanathan

Computational BPA Researcher at Autodesk and a Graduate Student Researcher at Carnegie Mellon University. Researching on the confluence of computational methods, Data Science and Artificial Intelligence for fostering synergies through the coexistence of early-phase design and real-time impact assessment.

Spyridon Ampanavos

Doctoral candidate at Harvard Graduate School of Design, researching the use of machine learning methods to support early phase performance driven design. Background in Architectural Engineering and Computational Design.

What will I learn?

LEARNING OBJECTIVE 1

Learn how to design workflows with Generative Design in Revit and Dynamo for building synthetic data sets to be used in training machine-learning models.

LEARNING OBJECTIVE 2

Discover the diversity of mass model geometry required to represent a comprehensive set of possible building types.

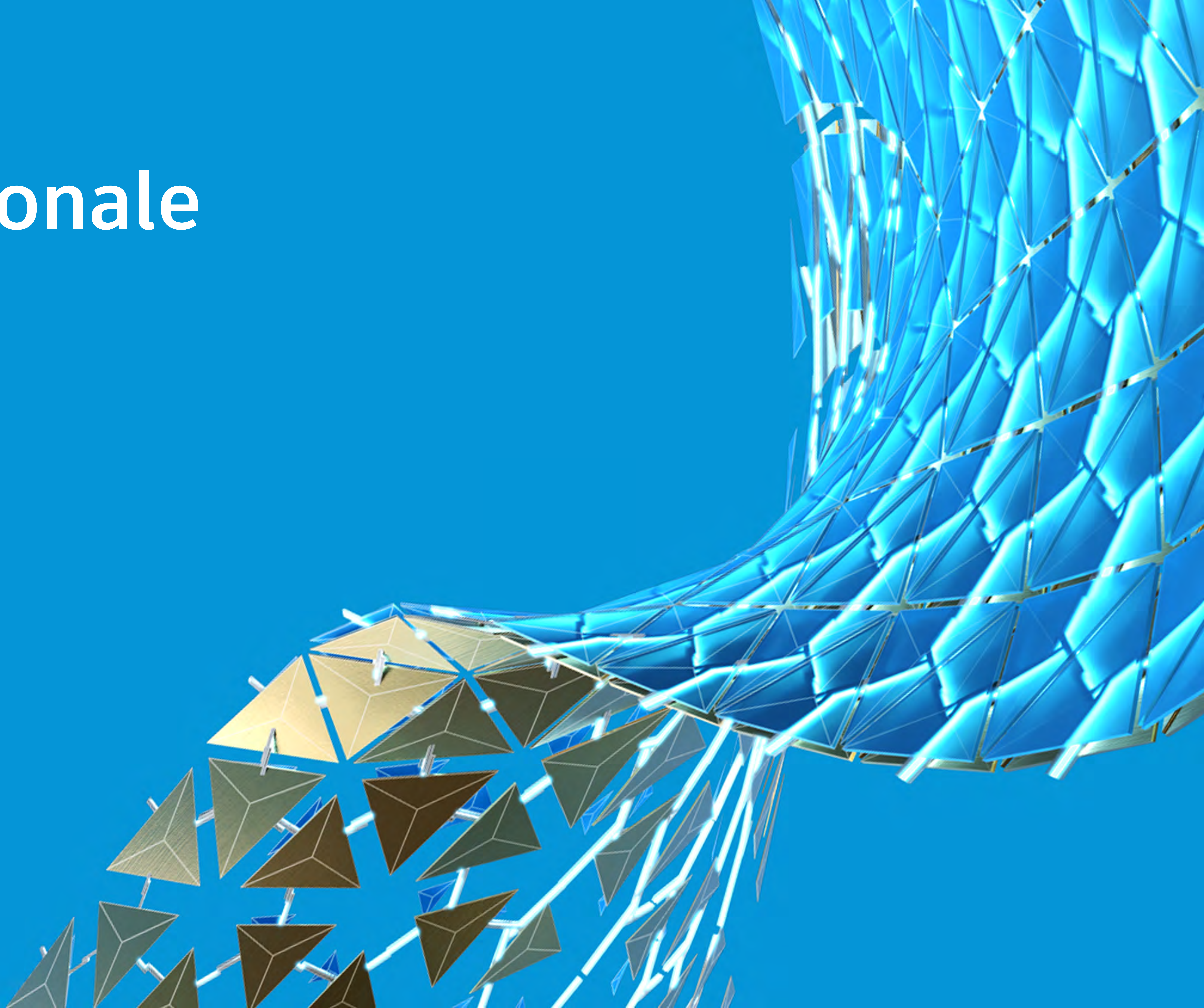
LEARNING OBJECTIVE 3

Learn how to represent your data to be used in training machine-learning models.

LEARNING OBJECTIVE 4

Discover potential uses of machine-learning models toward achieving faster analysis in early conceptual design stages.

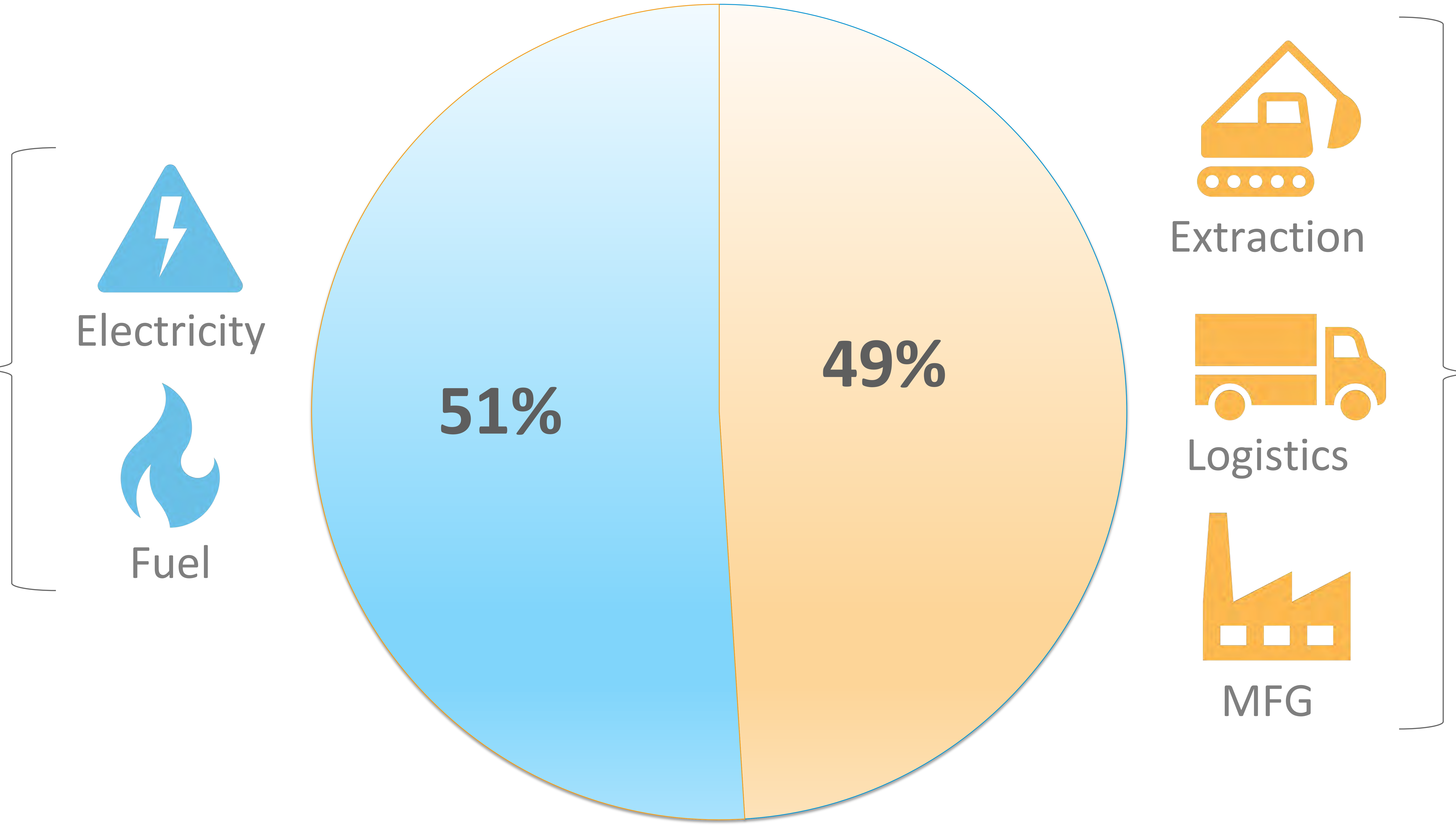
Project Rationale



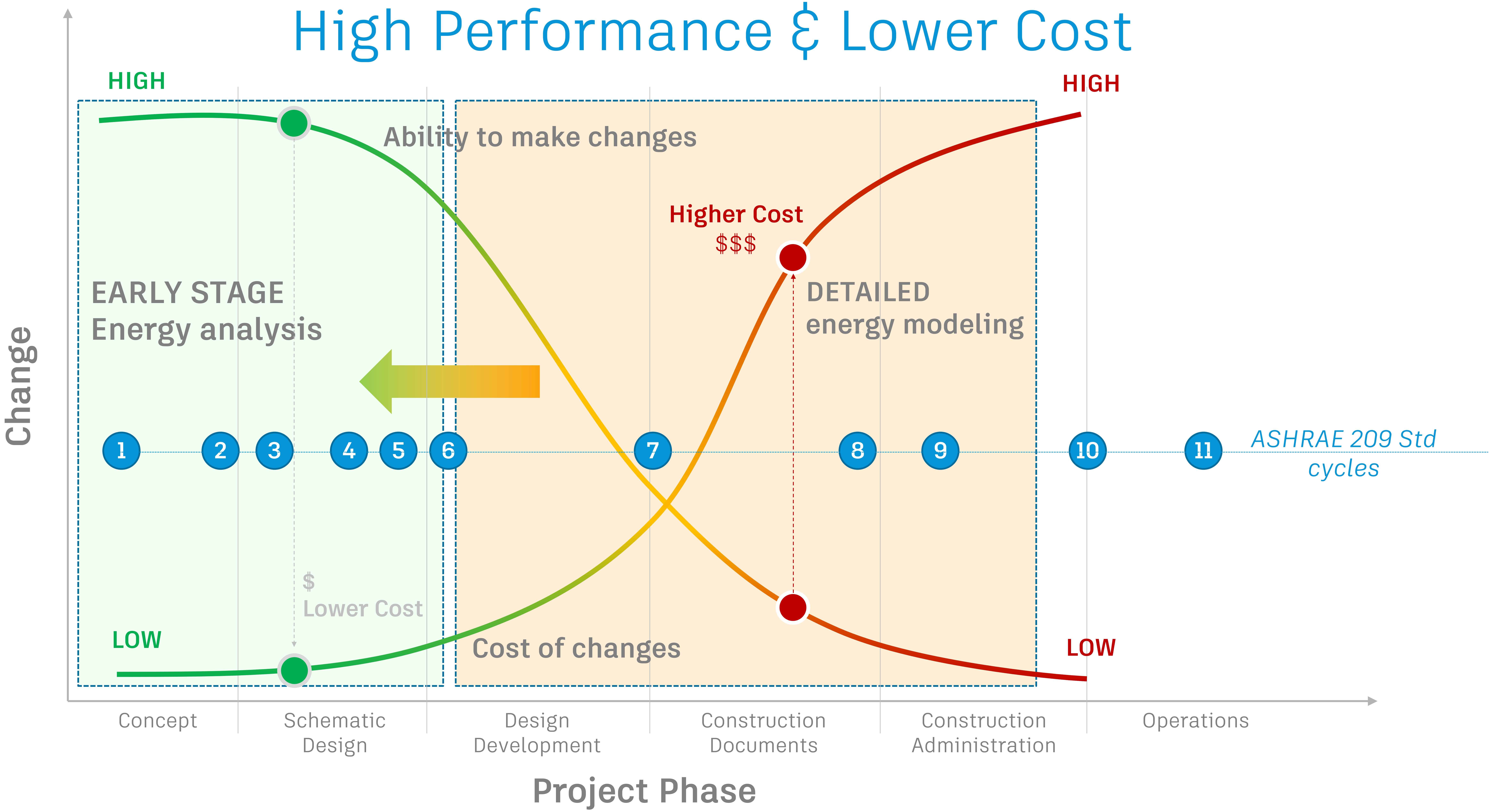

40%
CARBON EMISSIONS

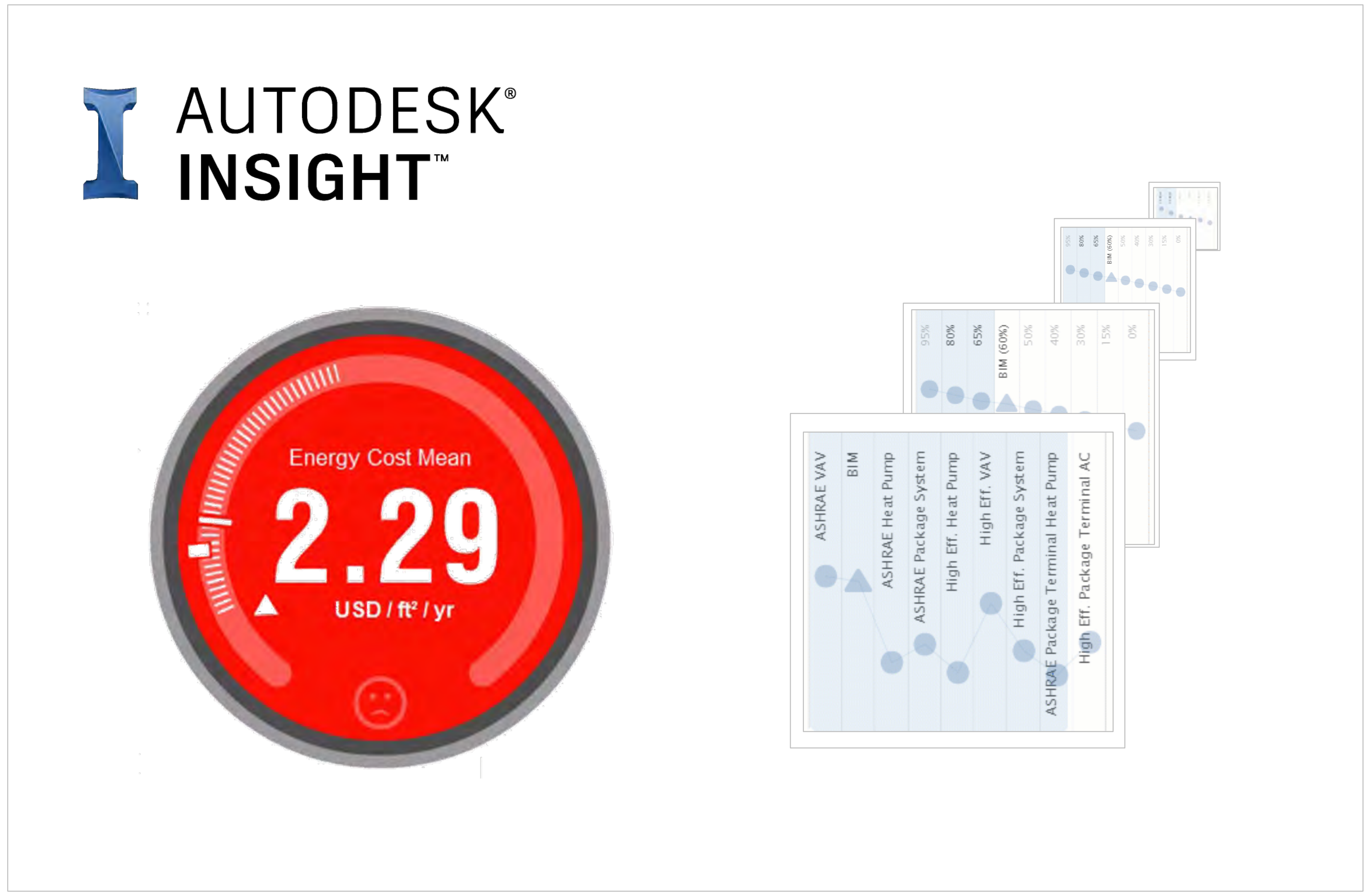

↓50%
BY 2030

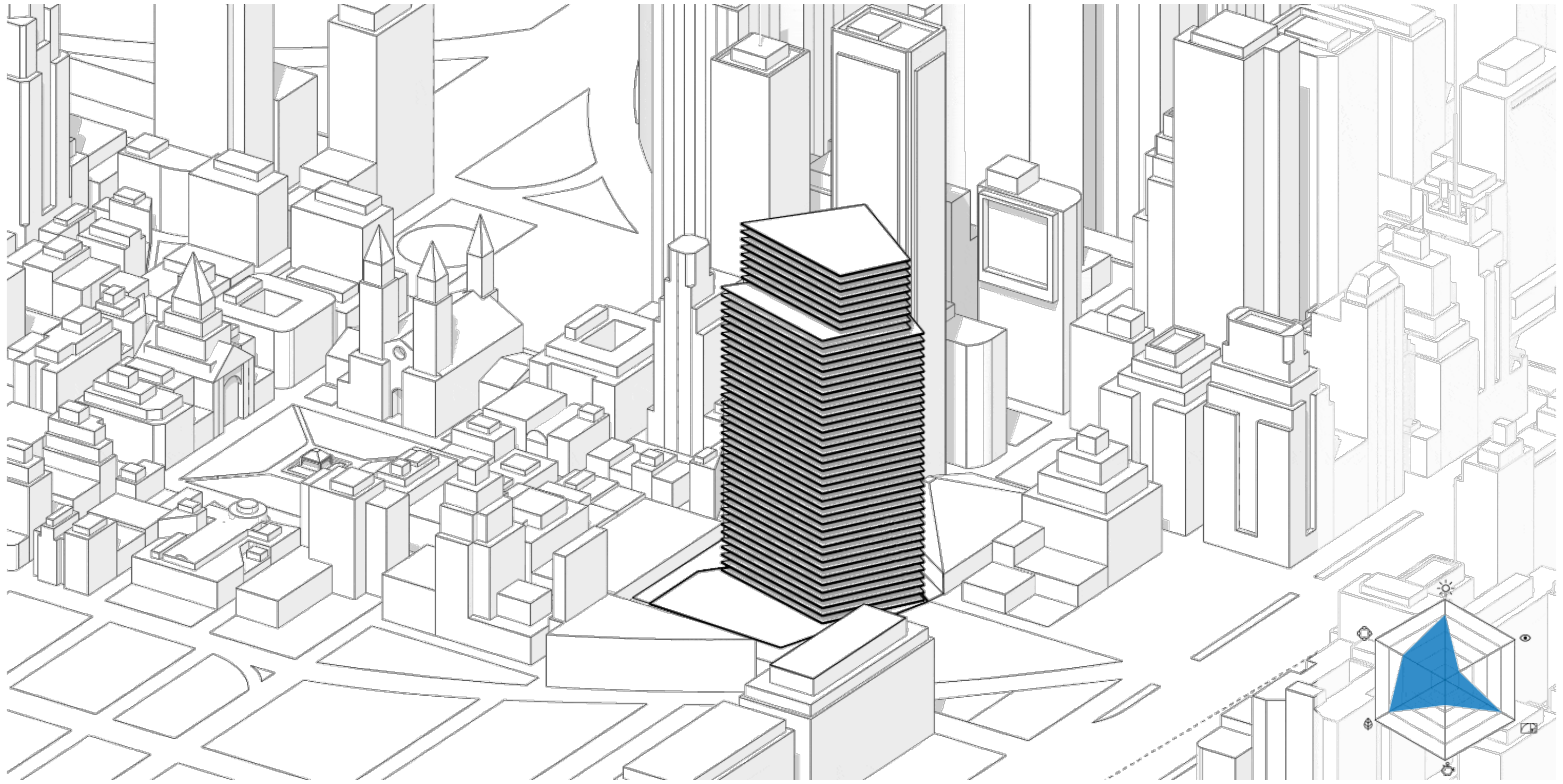
Total carbon from buildings → 2050



Early Stage Analysis =
High Performance & Lower Cost



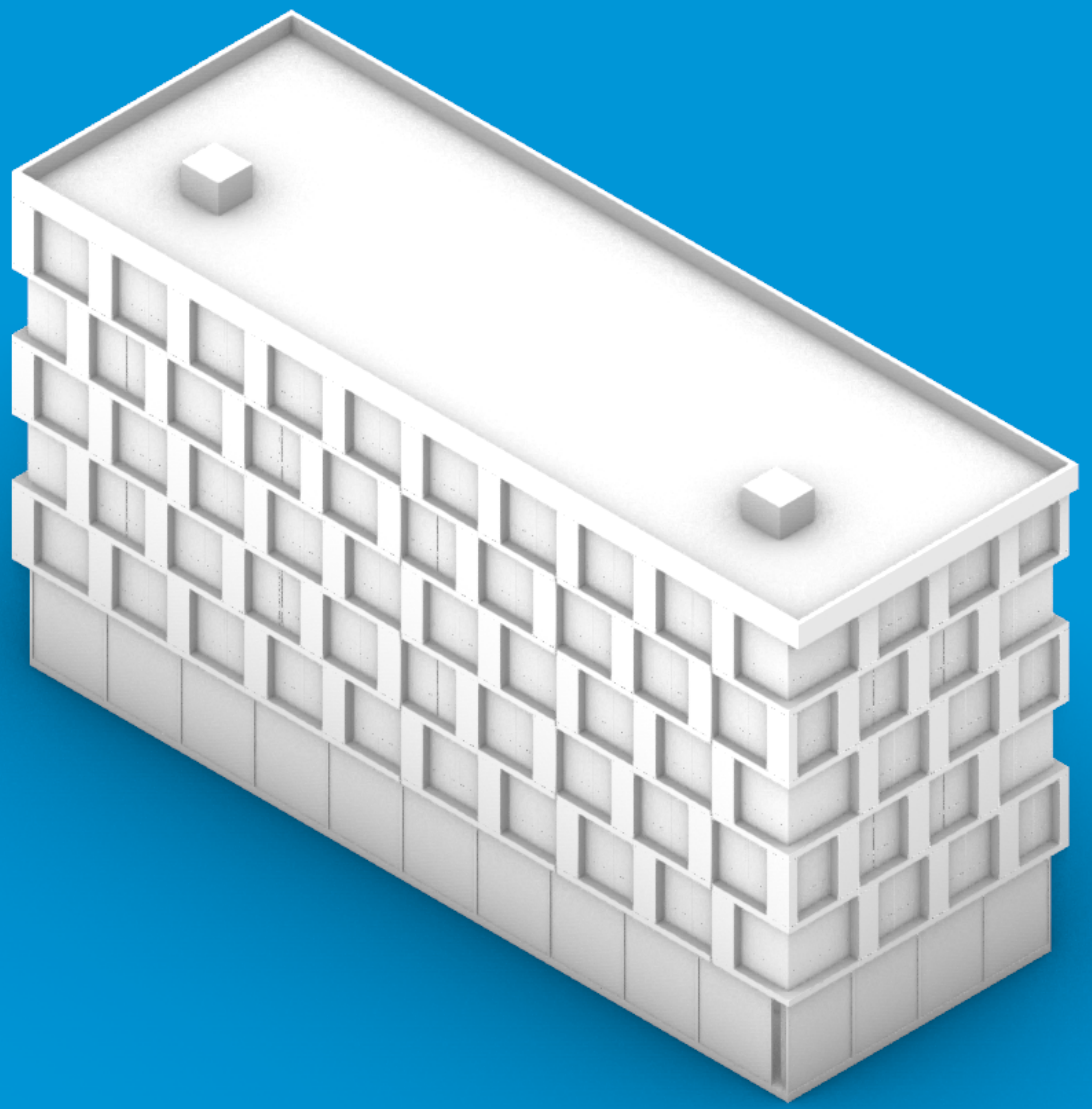

AUTODESK®
INSIGHT™
Energy Cost Mean
2.29
USD / ft² / yr
ASHRAE VAV
BIM
ASHRAE Heat Pump
ASHRAE Package System
High Eff. Heat Pump
High Eff. VAV
High Eff. Package System
ASHRAE Package Terminal Heat Pump
High Eff. Package Terminal AC

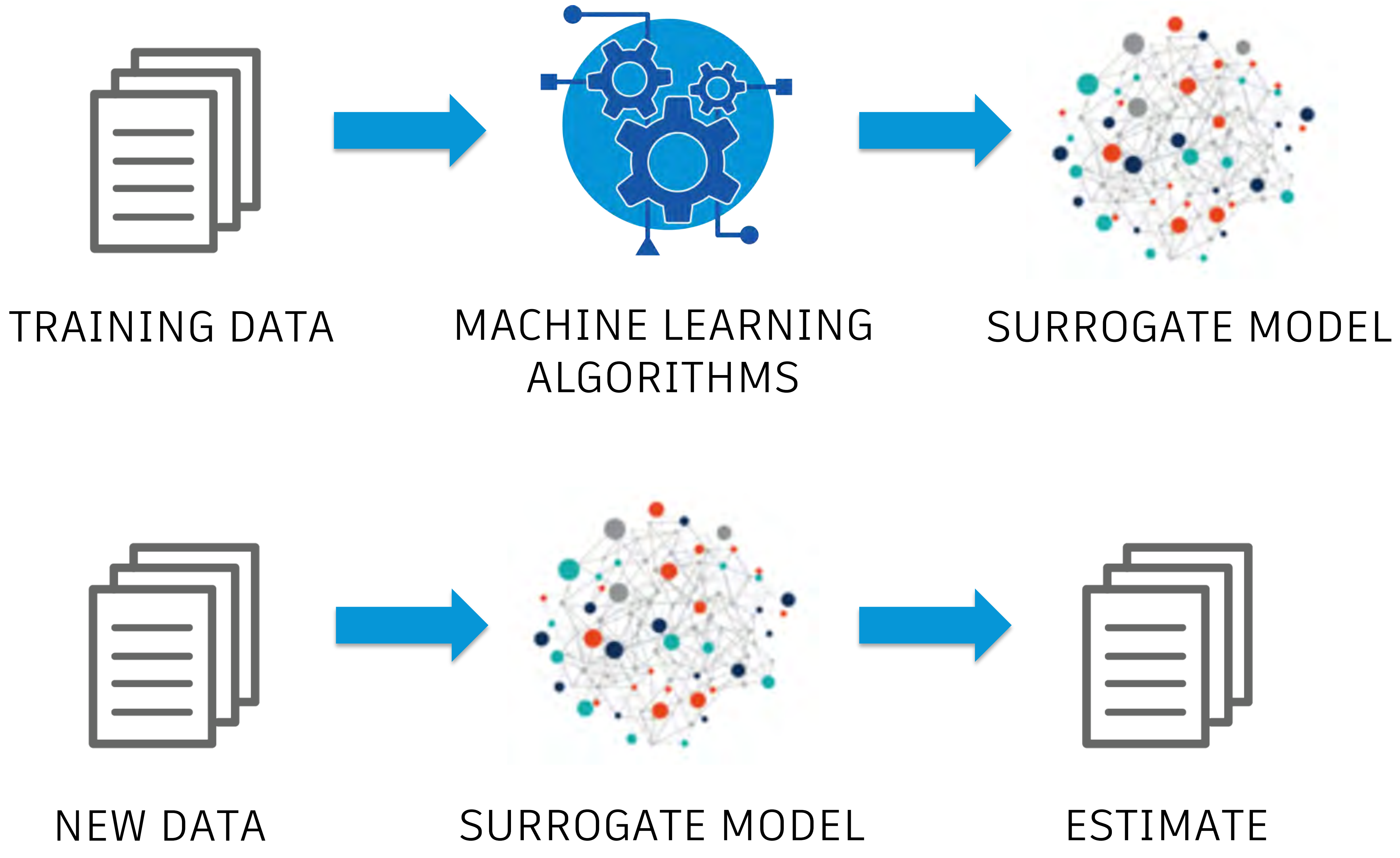


Energy Use Intensity (kWh/m2) :
201.42

Energy Cost Mean ($/m2) :
3021.3

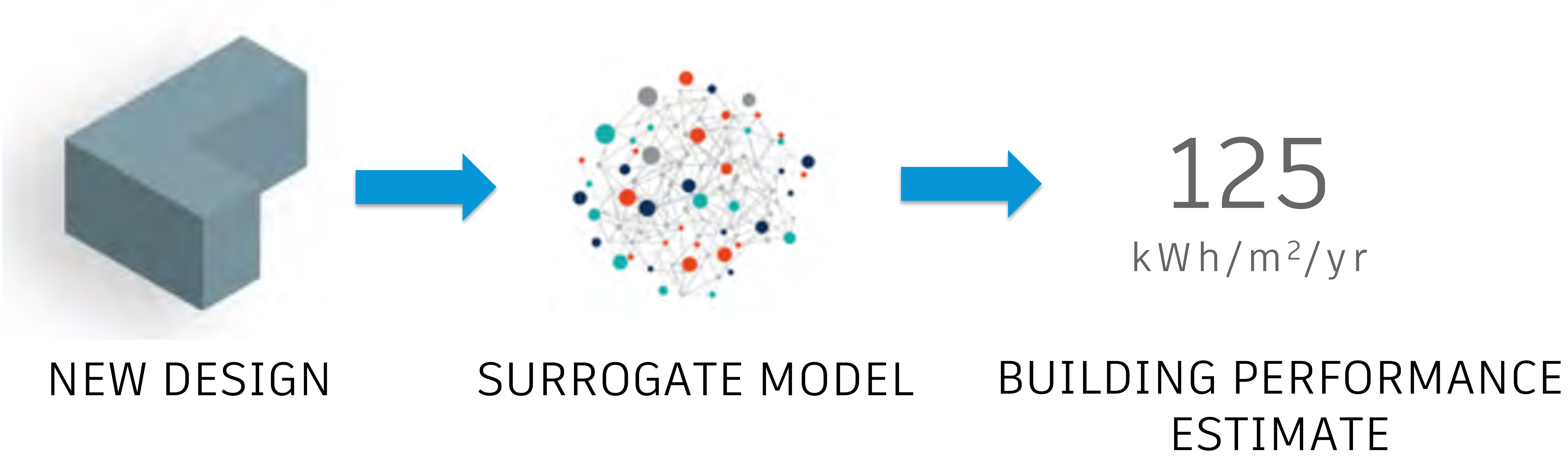
Embodied Carbon (kg-CO2e/m2) :
273.9312





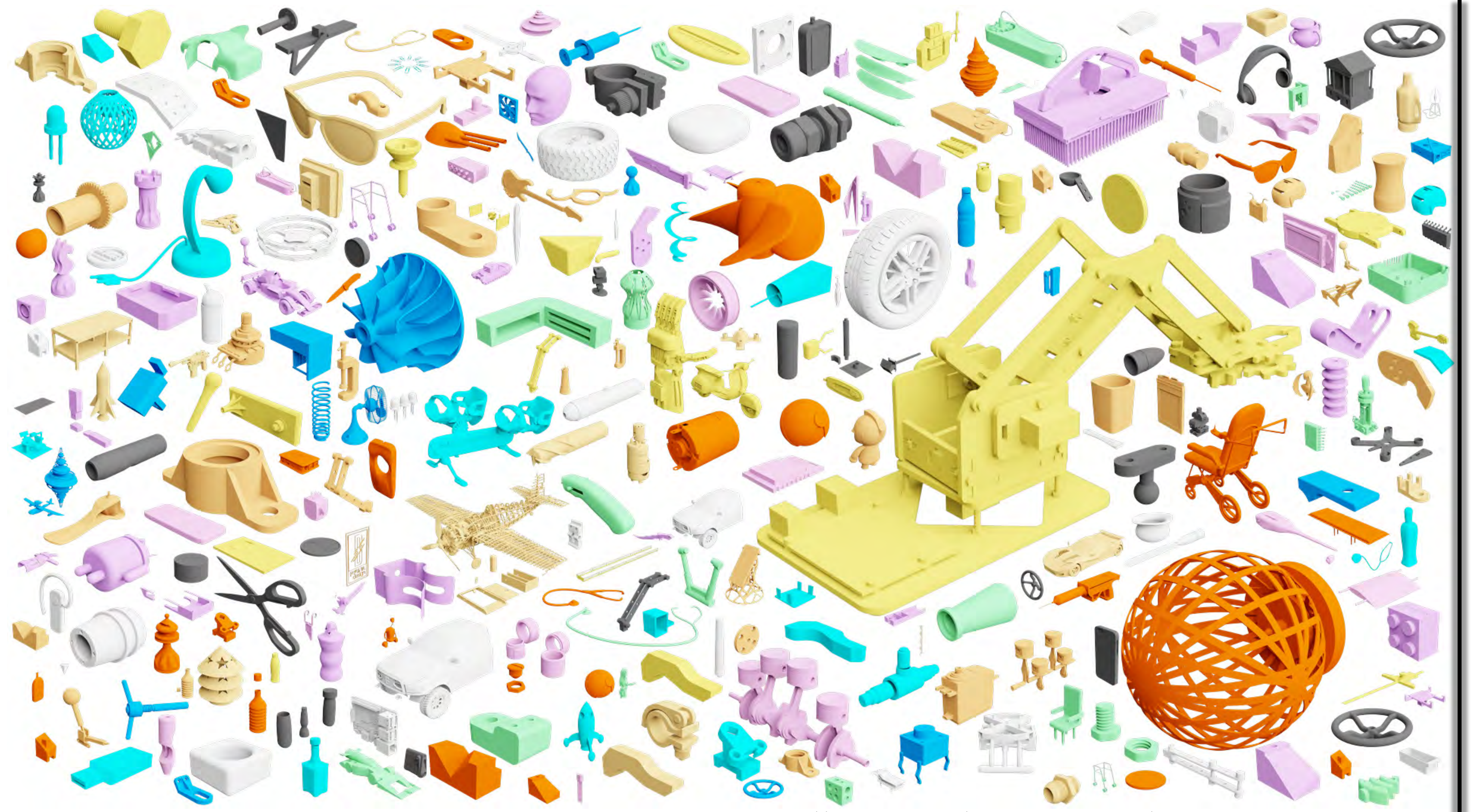
Surrogate modeling

Statistical models can be used as surrogates of detailed simulation models.

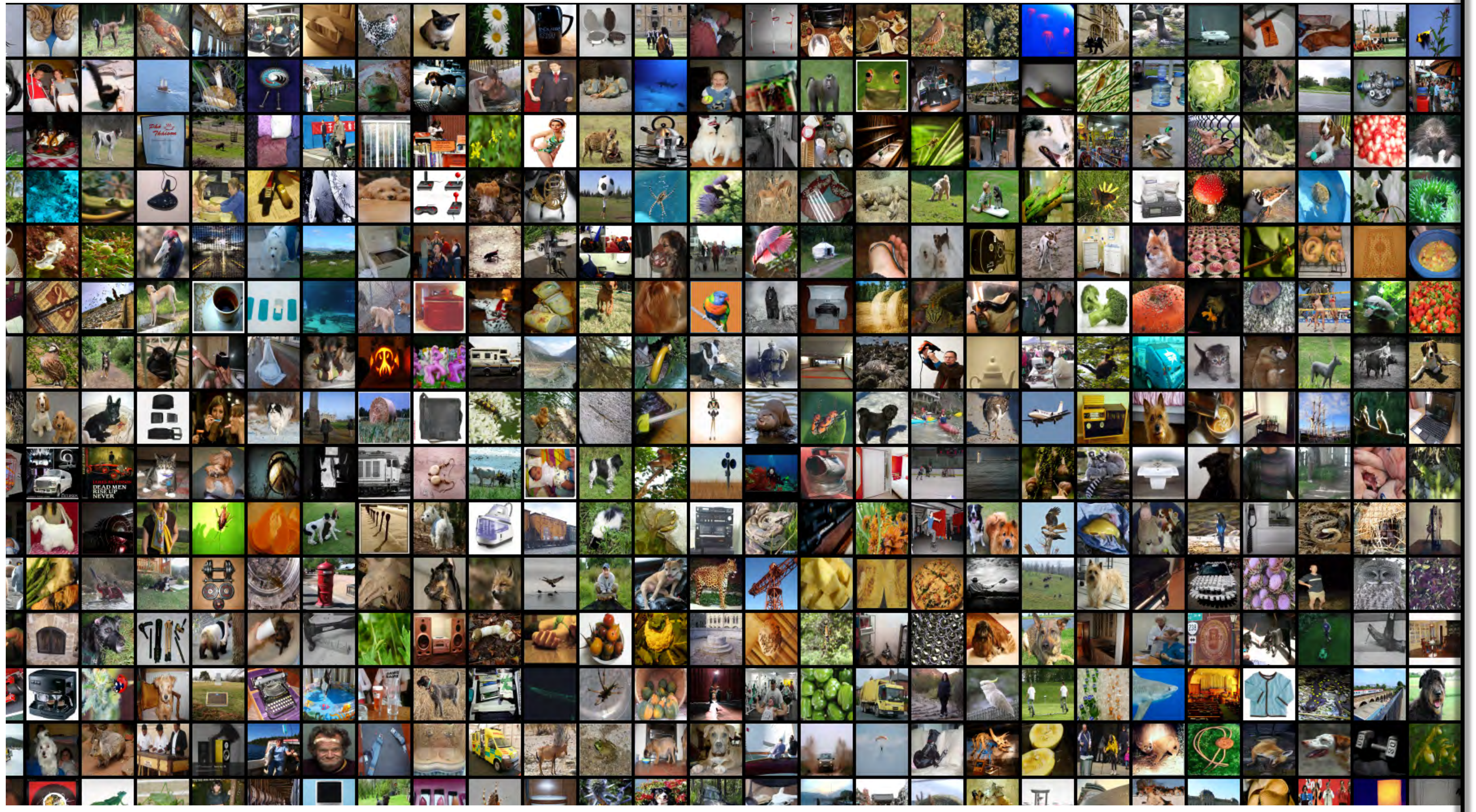


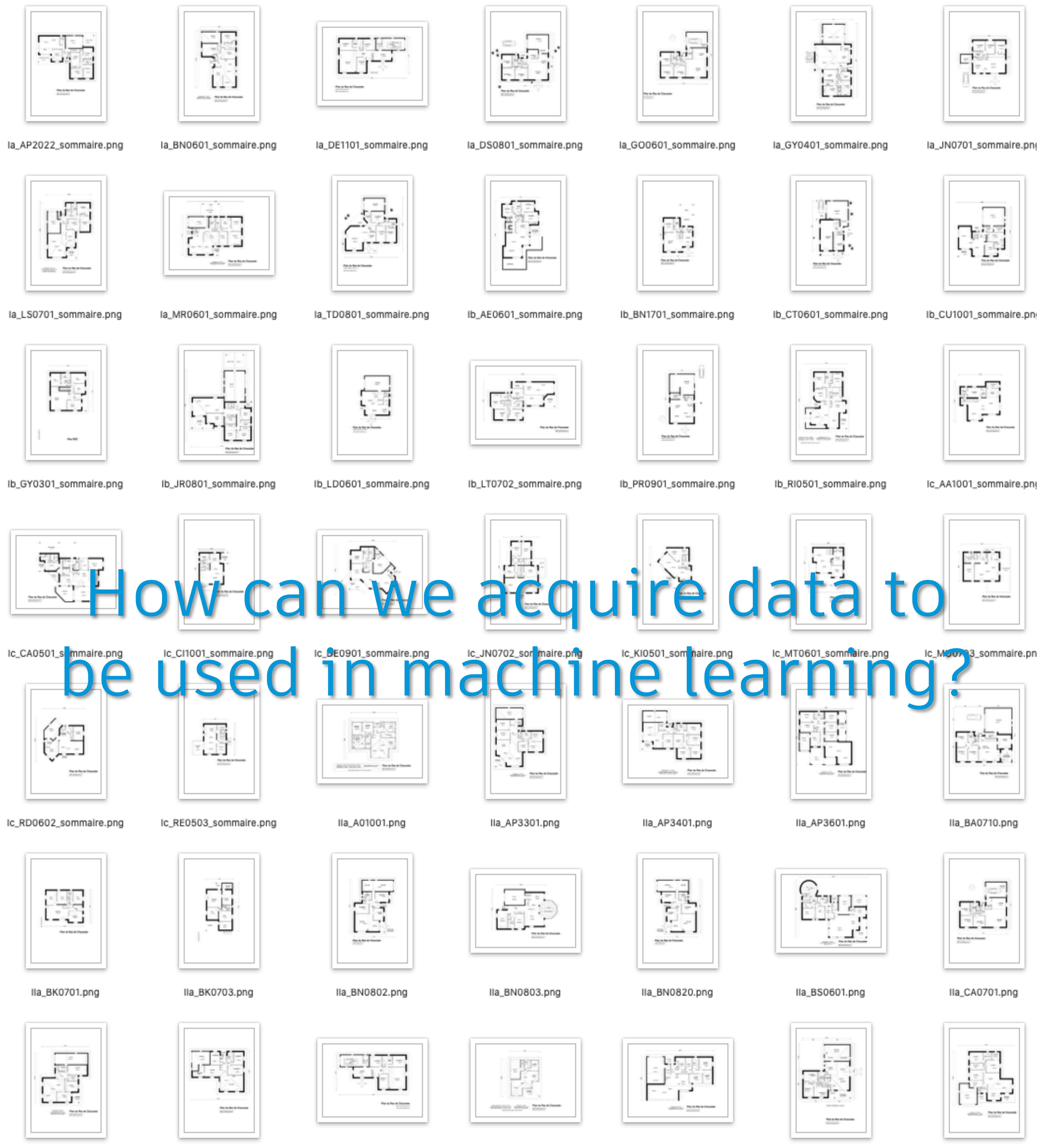
How can a surrogate model help in the conceptual design stage?

Surrogate models can be evaluated instantly and provide estimates of the building performance. This enables designers to rapidly assess a design concept and explore the design space.



Fusion 360 Gallery Dataset - https://github.com/AutodeskAILab/Fusion360GalleryDataset

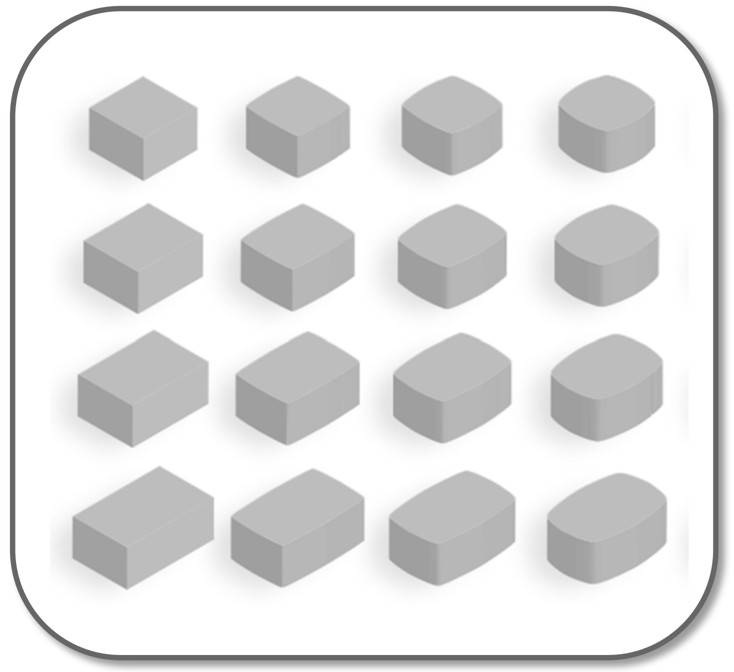




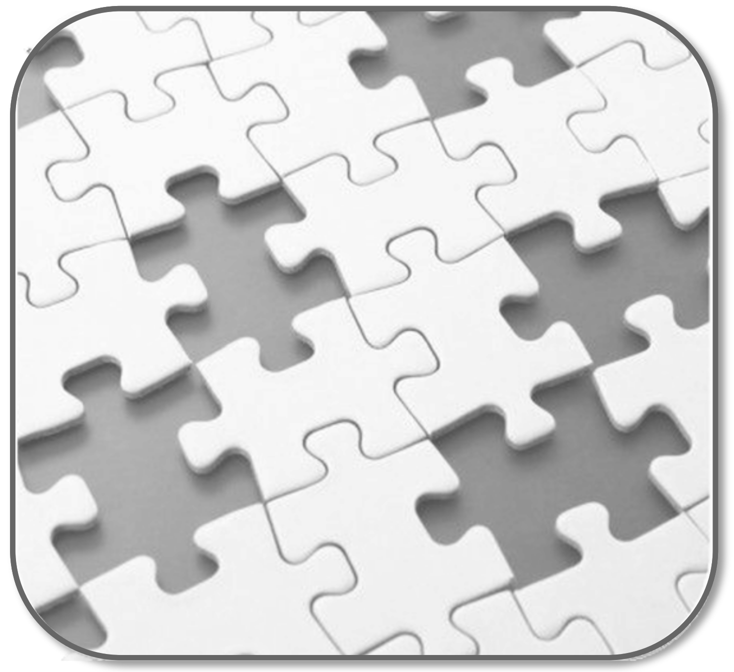
How can we acquire data to be used in machine learning?

Floor plans from the CVC-FP dataset - http://dag.cvc.uab.es/resources/floorplans/

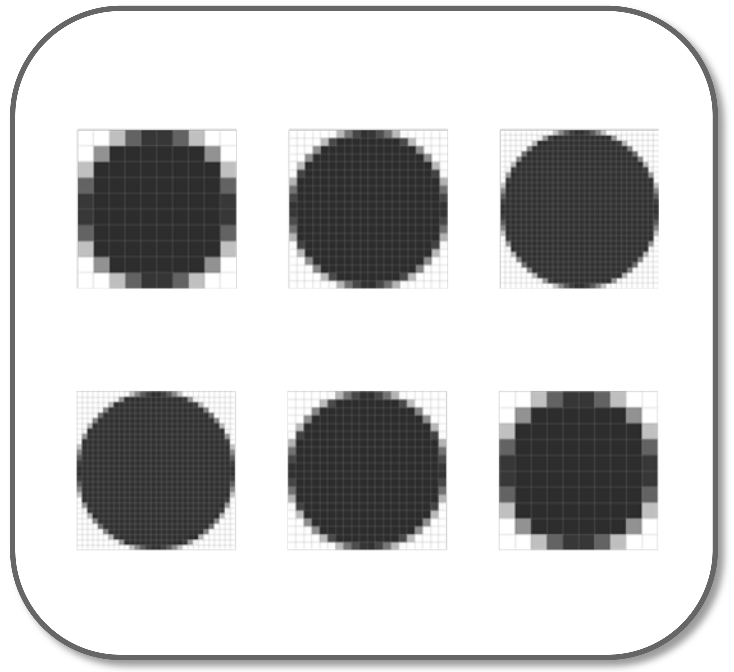
NOT ENOUGH



NOT DIVERSE



INCOMPLETE



INCOSISTENT
QUALITY & FORMAT

Limitations of collected data in AEC

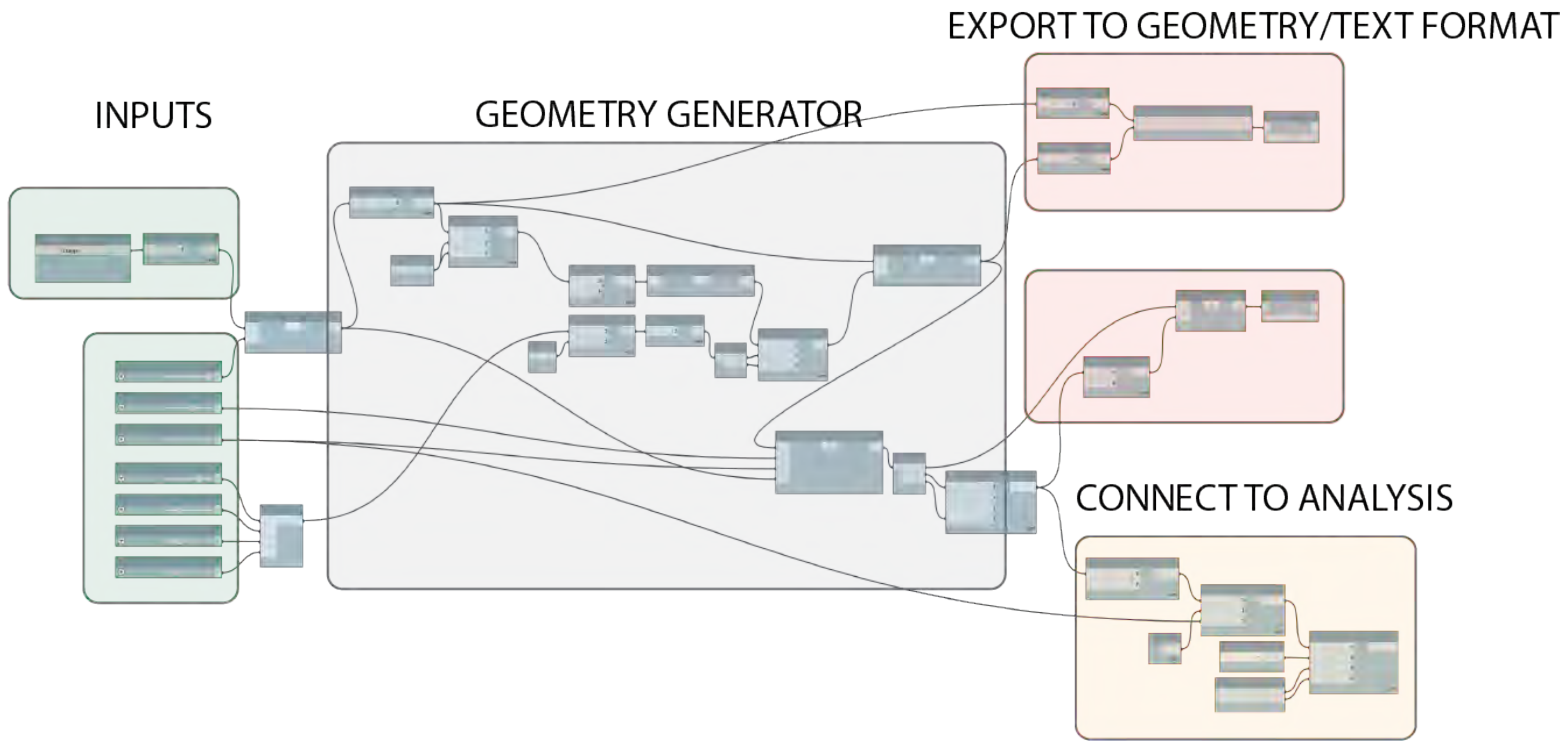
- LARGE SIZE
- GROUND TRUTH LABELS
- PARAMETRIC REPRESENTATION
- EXPANDABLE
- VARIATION
- BALANCED

Koch et al. (2019), ABC: A Big CAD Model Dataset For Geometric Deep Learning

Which are the properties of a good dataset?

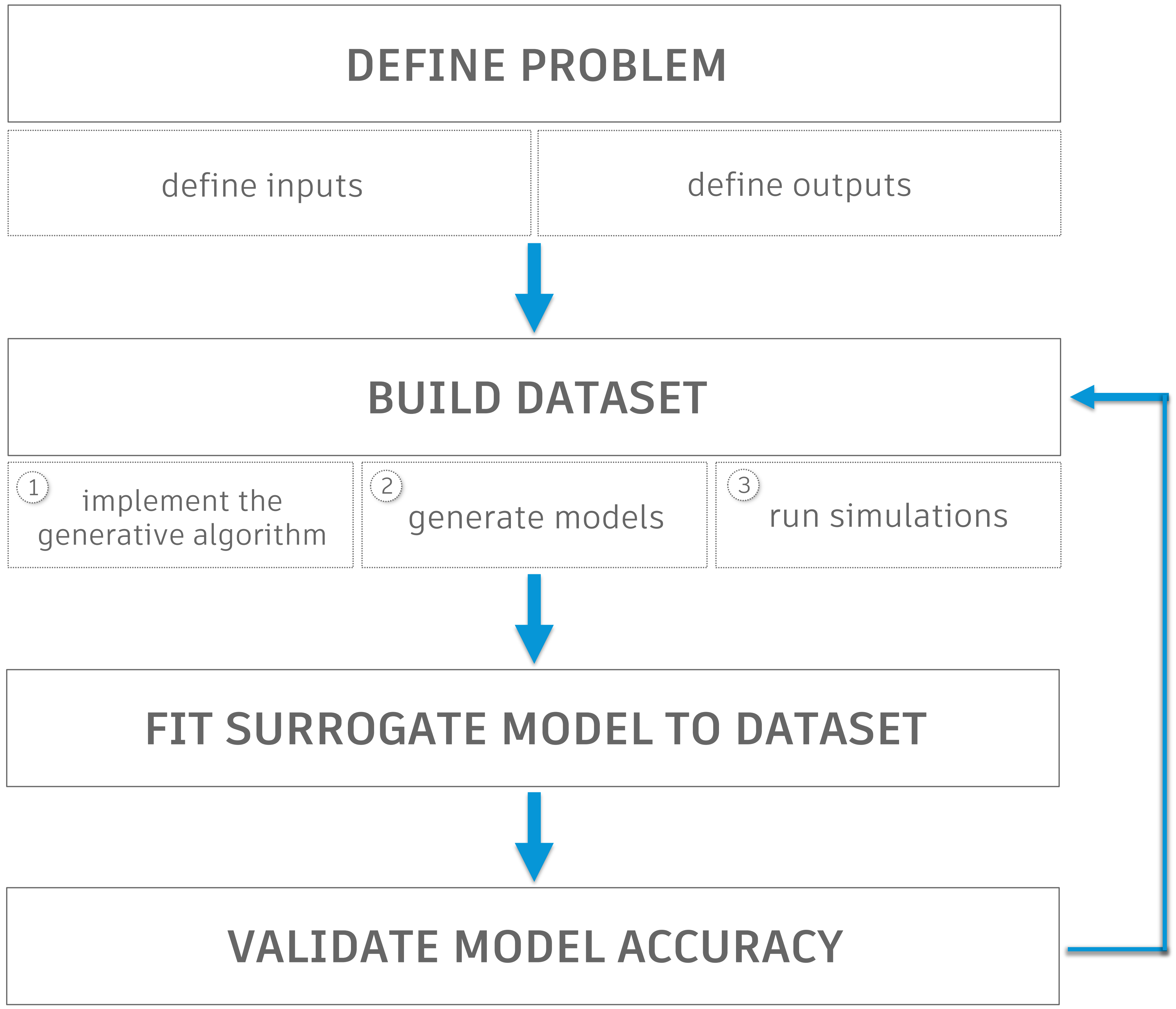
Synthetic data creation via generative design workflows

Dynamo

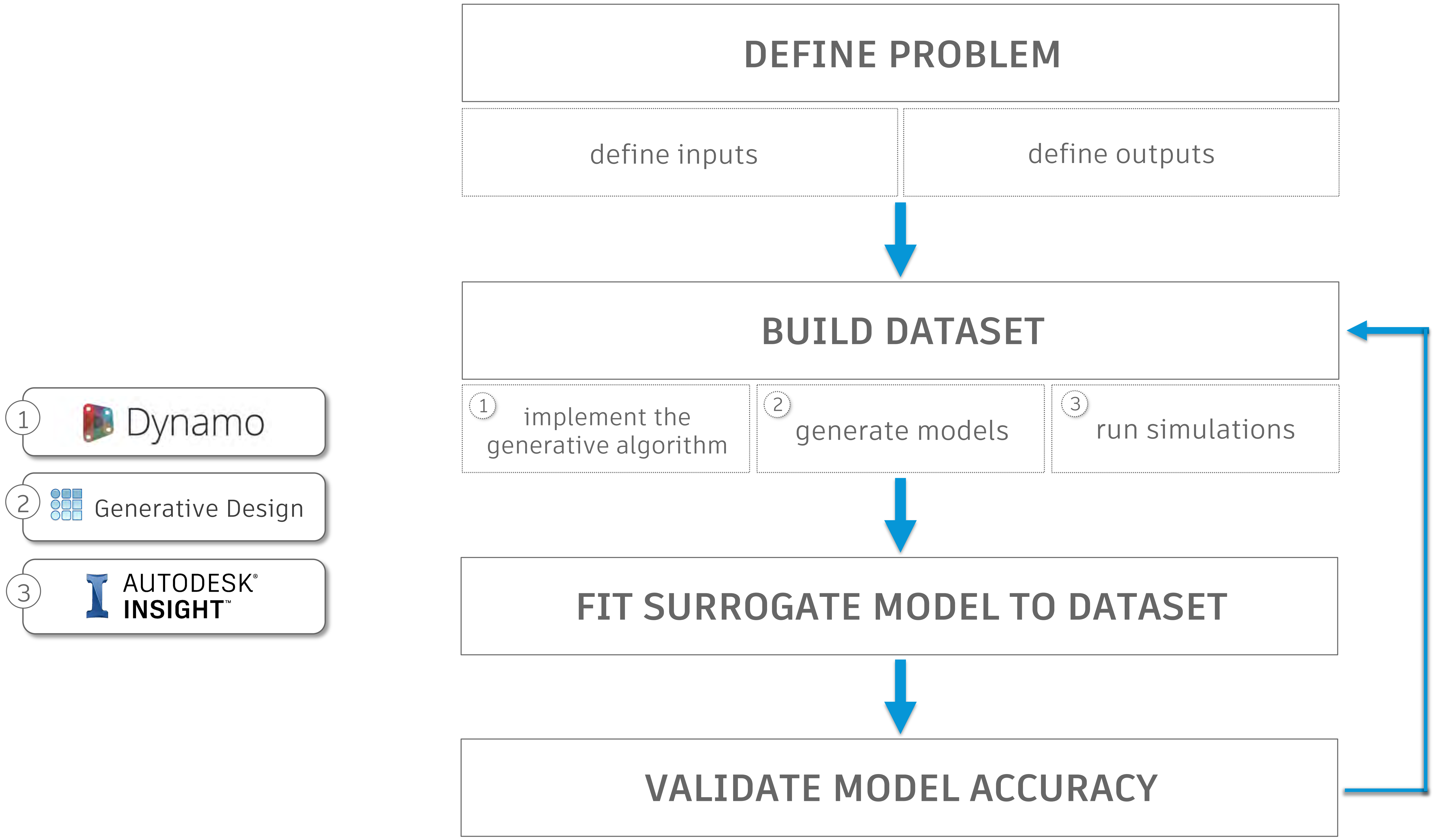


Generative Design in Revit

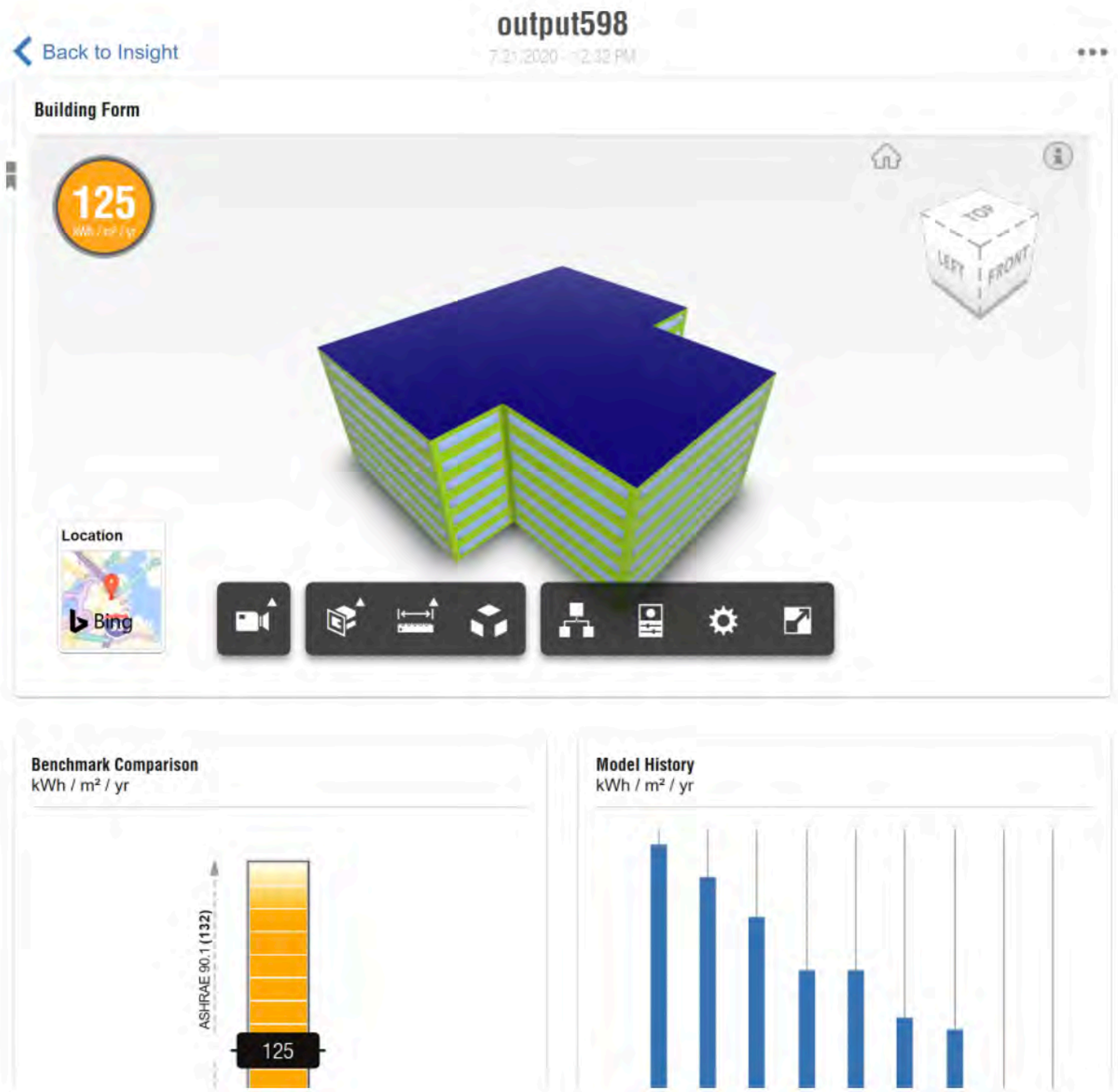
Generative Design Primer – https://www.generativedesign.org/



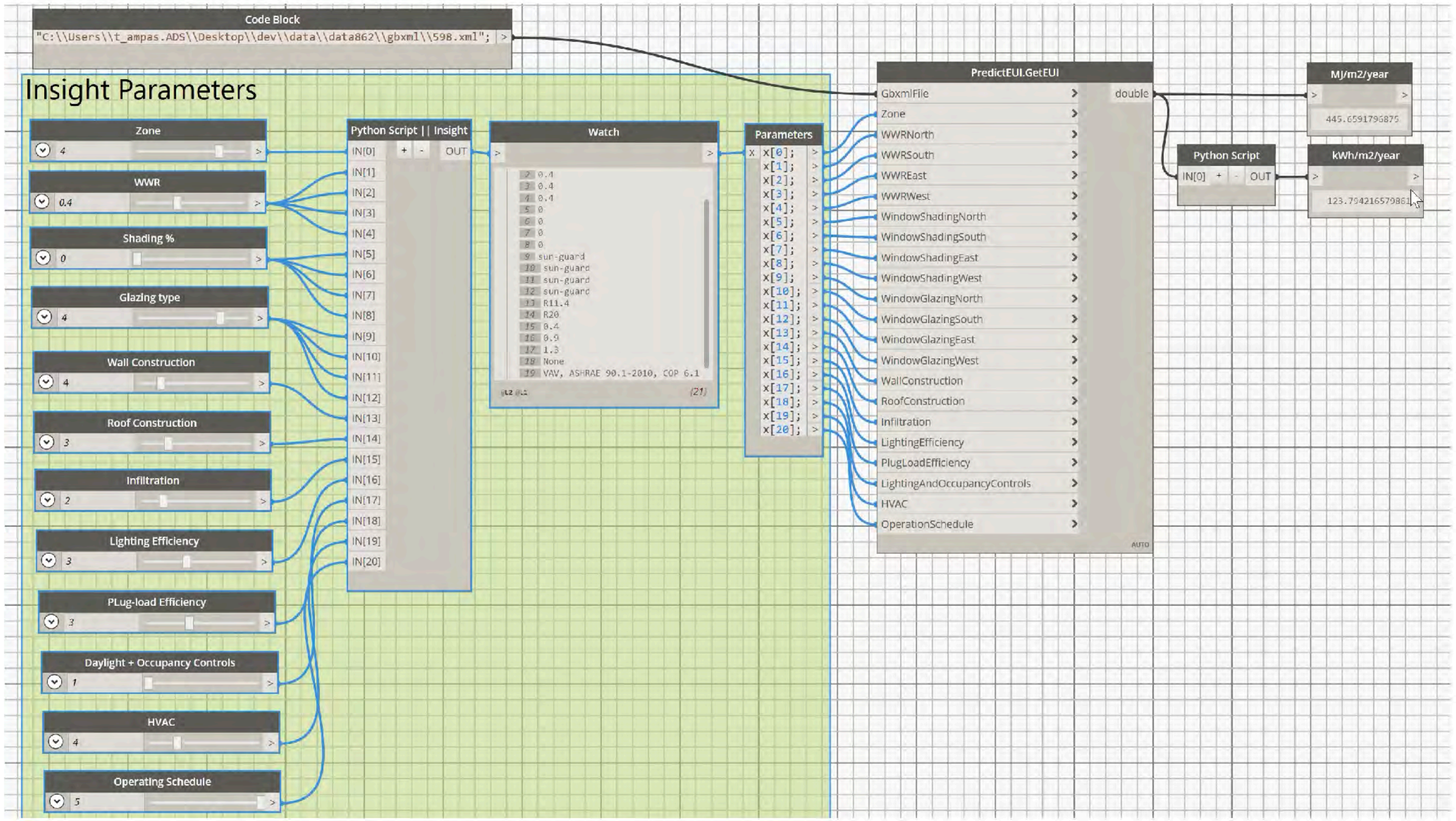
How do we structure a workflow for deriving a surrogate model?



How do we structure a workflow for deriving a surrogate model?



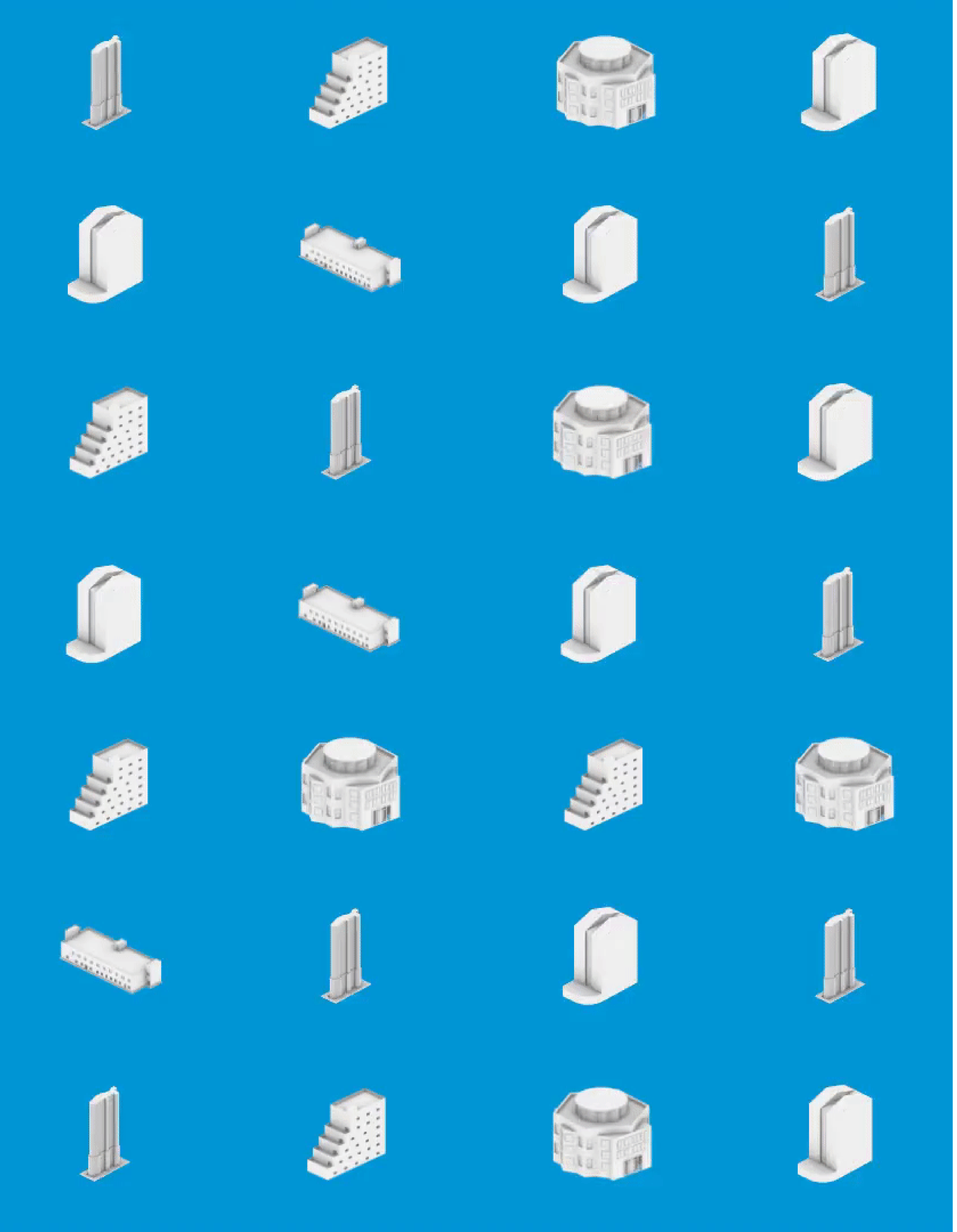
output598
Back to Insight
Building Form
125
Location
Bing
Benchmark Comparison
kWh / m² / yr
ASHRAE 90.1 (132)
125
Model History
kWh / m² / yr



Code Block
"C:\\Users\\t_ampas.ADS\\Desktop\\dev\\data\\data862\\gbxml\\598.xml";
Insight Parameters
Zone
4
WWR
0.4
Shading %
0
Glazing type
4
Wall Construction
4
Roof Construction
3
Infiltration
2
Lighting Efficiency
3
PLug-load Efficiency
3
Daylight + Occupancy Controls
1
HVAC
4
Operating Schedule
5
Python Script || Insight
Watch
Parameters
PredictEUI.GetEUI
GbxmlFile
Zone
WWRNorth
WWRSouth
WWREast
WWRWest
WindowShadingNorth
WindowShadingSouth
WindowShadingEast
WindowShadingWest
WindowGlazingNorth
WindowGlazingSouth
WindowGlazingEast
WindowGlazingWest
WallConstruction
RoofConstruction
Infiltration
LightingEfficiency
PlugLoadEfficiency
LightingAndOccupancyControls
HVAC
OperationSchedule
double
Python Script
MJ/m2/year
445.6591796875
kWh/m2/year
123.794216579861

Defining the problem for synthetic data generation

..the generative design way!



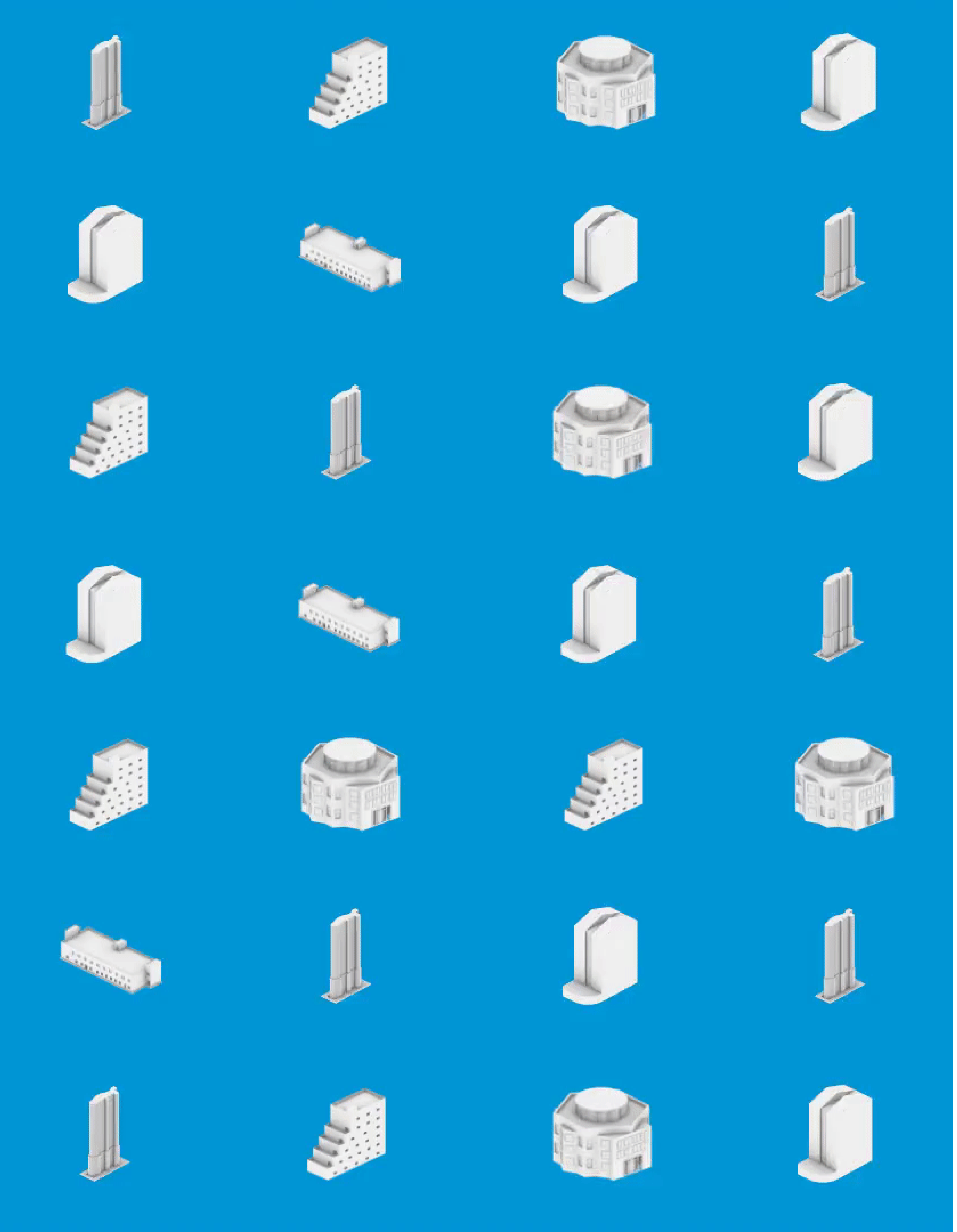
LARGE SIZE

GROUND TRUTH LABELS

PARAMETRIC REPRESENTATION

EXPANDABLE

VARIATION



LARGE SIZE

AUTOMATED GENERATION

GROUND TRUTH LABELS

SIMULATED PERFORMANCE RESULTS

PARAMETRIC REPRESENTATION

REPRODUCIBLE WORKFLOW

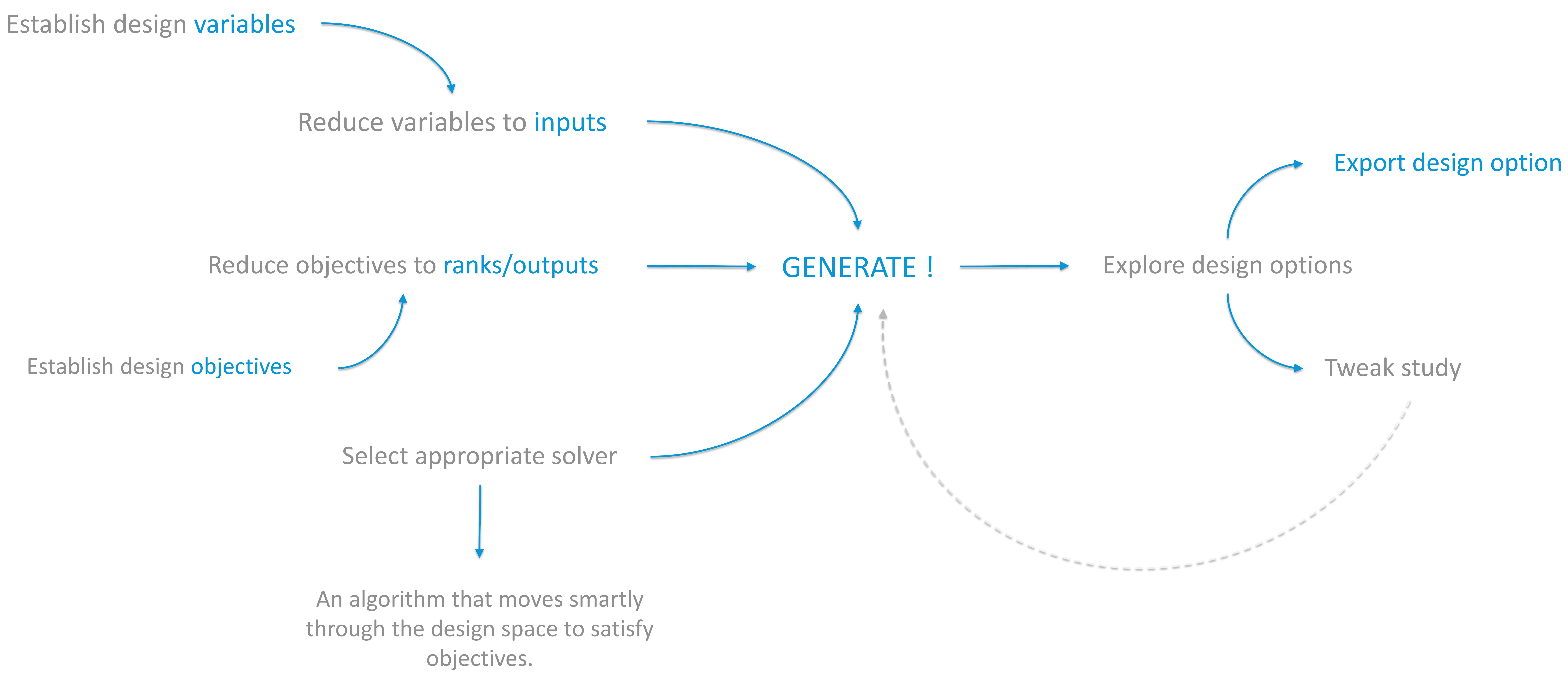
EXPANDABLE

SCALABLE WORKFLOW

VARIATION

UNBIASED

Generative design workflow

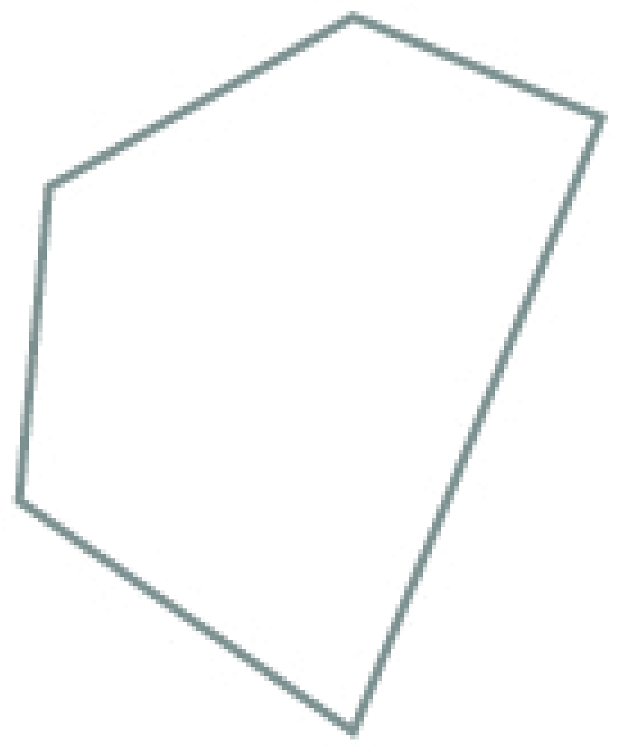


Generative design workflow

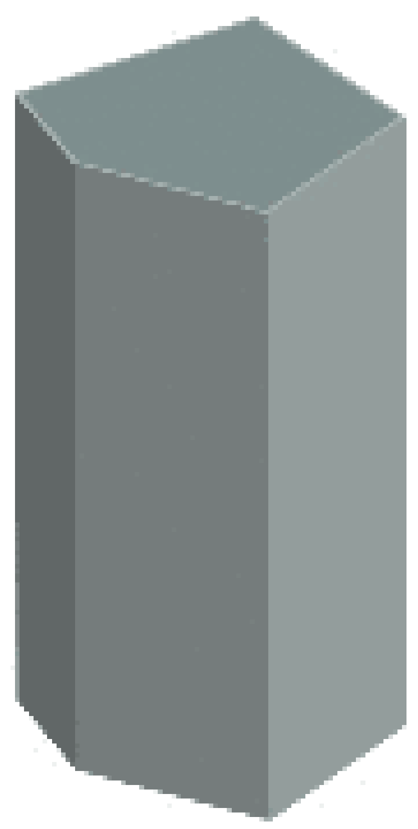
IDENTIFY

TARGET TAXONOMIES

Footprint level

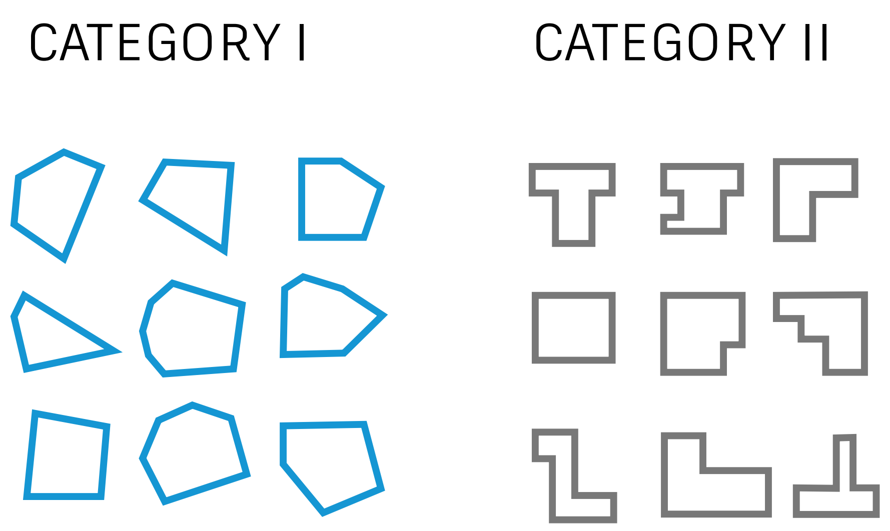


Massing level

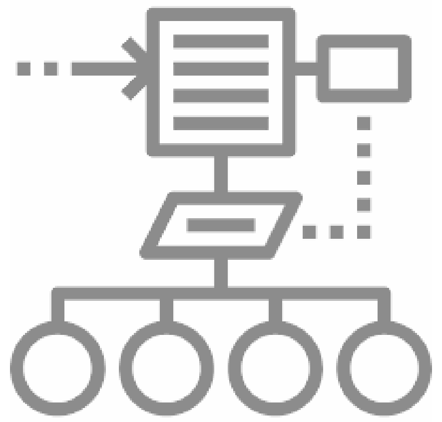


IDENTIFY

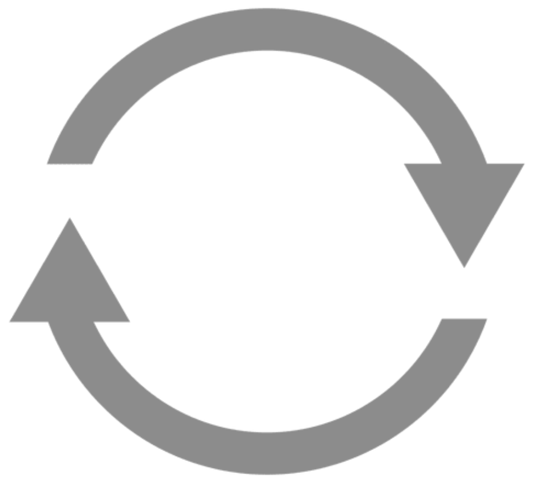
TARGET TAXONOMIES



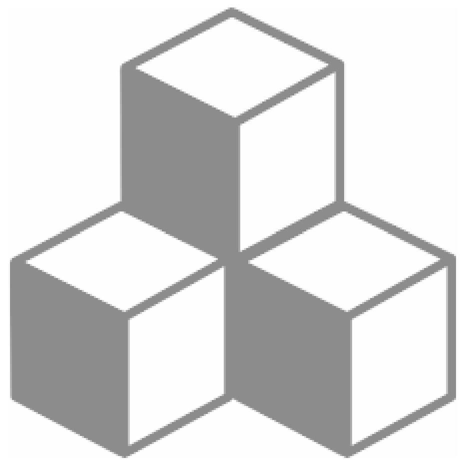
Categorically differentiable



Algorithmically representable



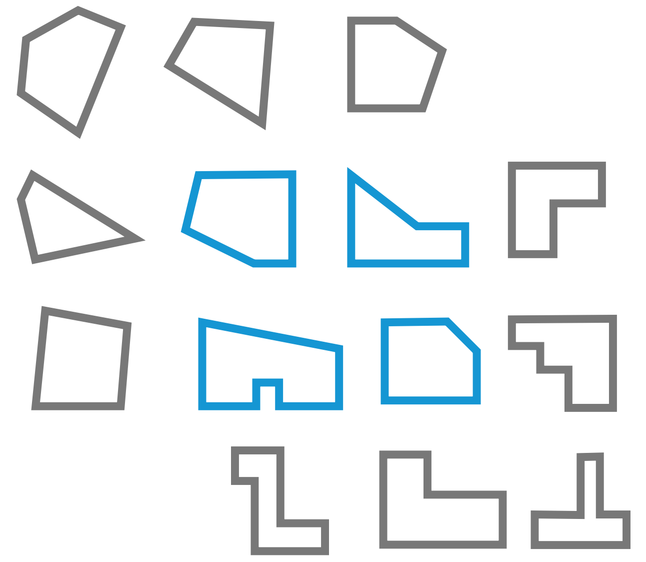
Reproducible and scalable



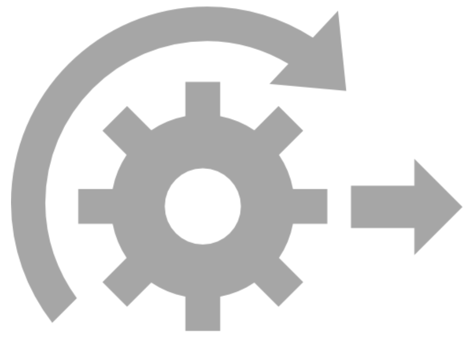
Modular generative workflows

IDENTIFY

TARGET TAXONOMIES

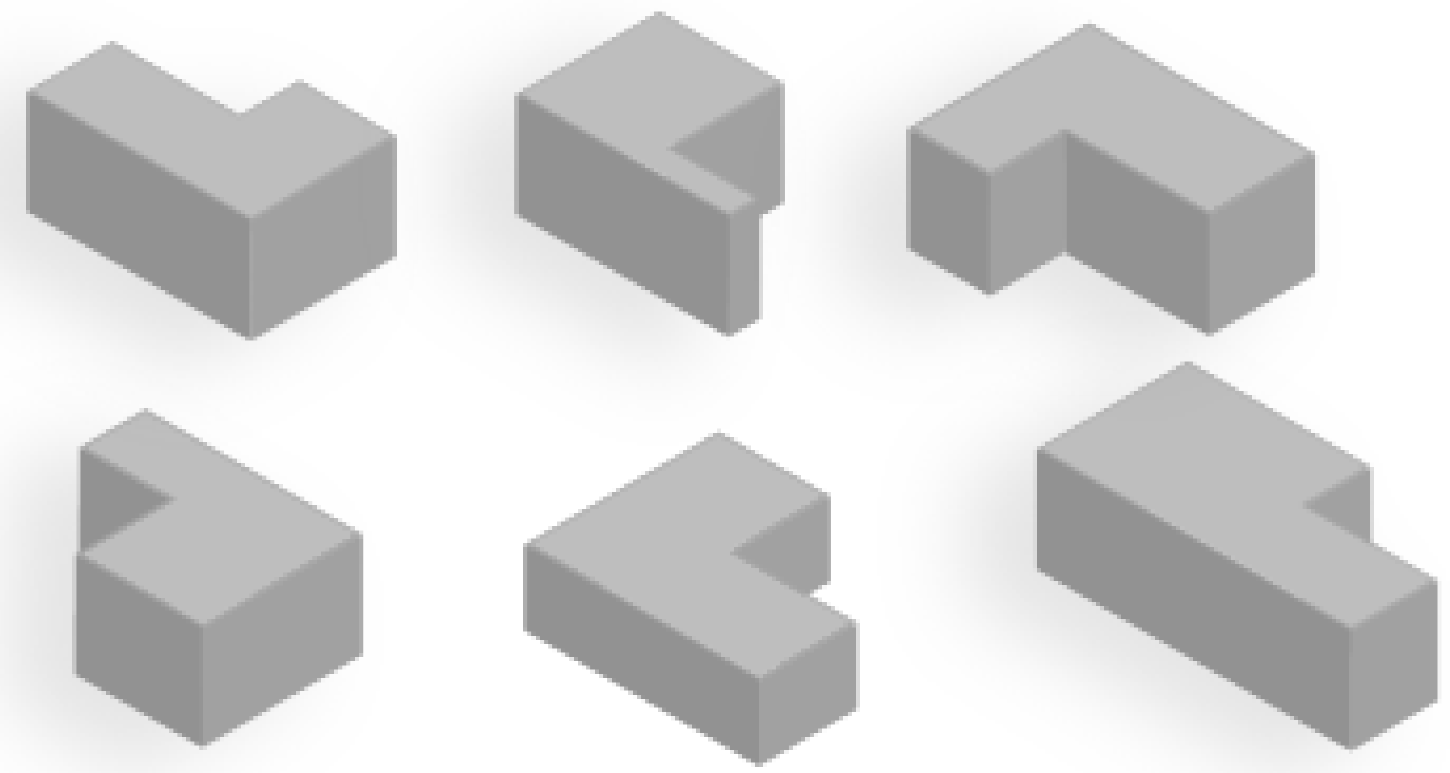


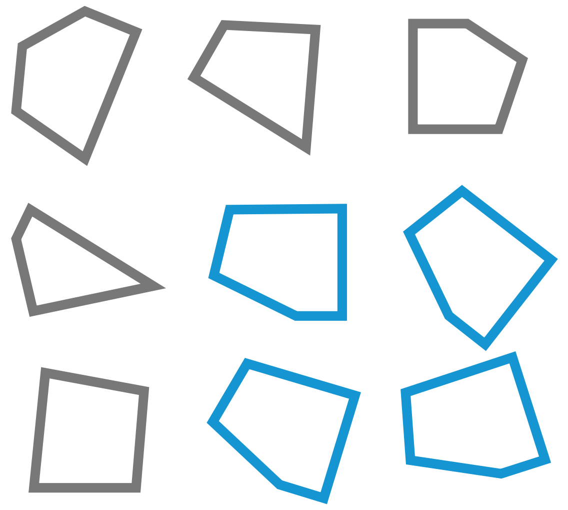
Avoid overlaps



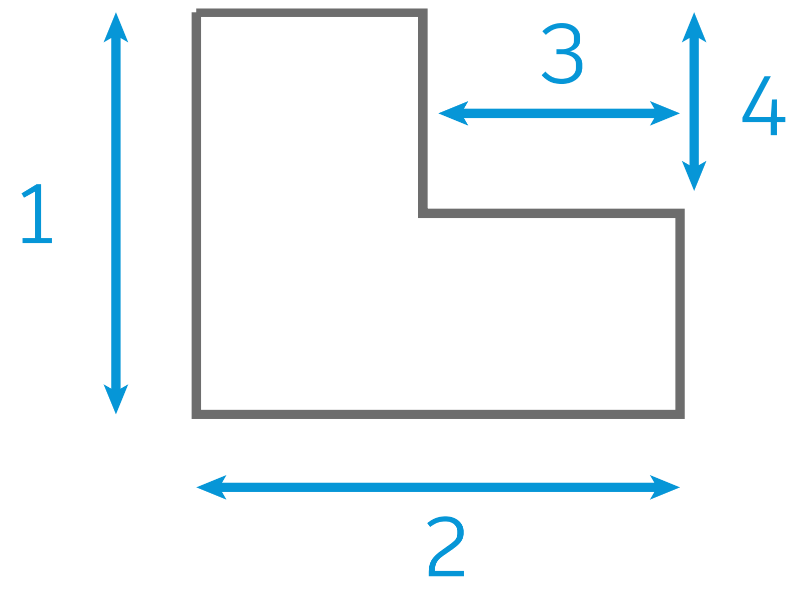
Breakdown algorithmically

Category of 'L' shaped geometries





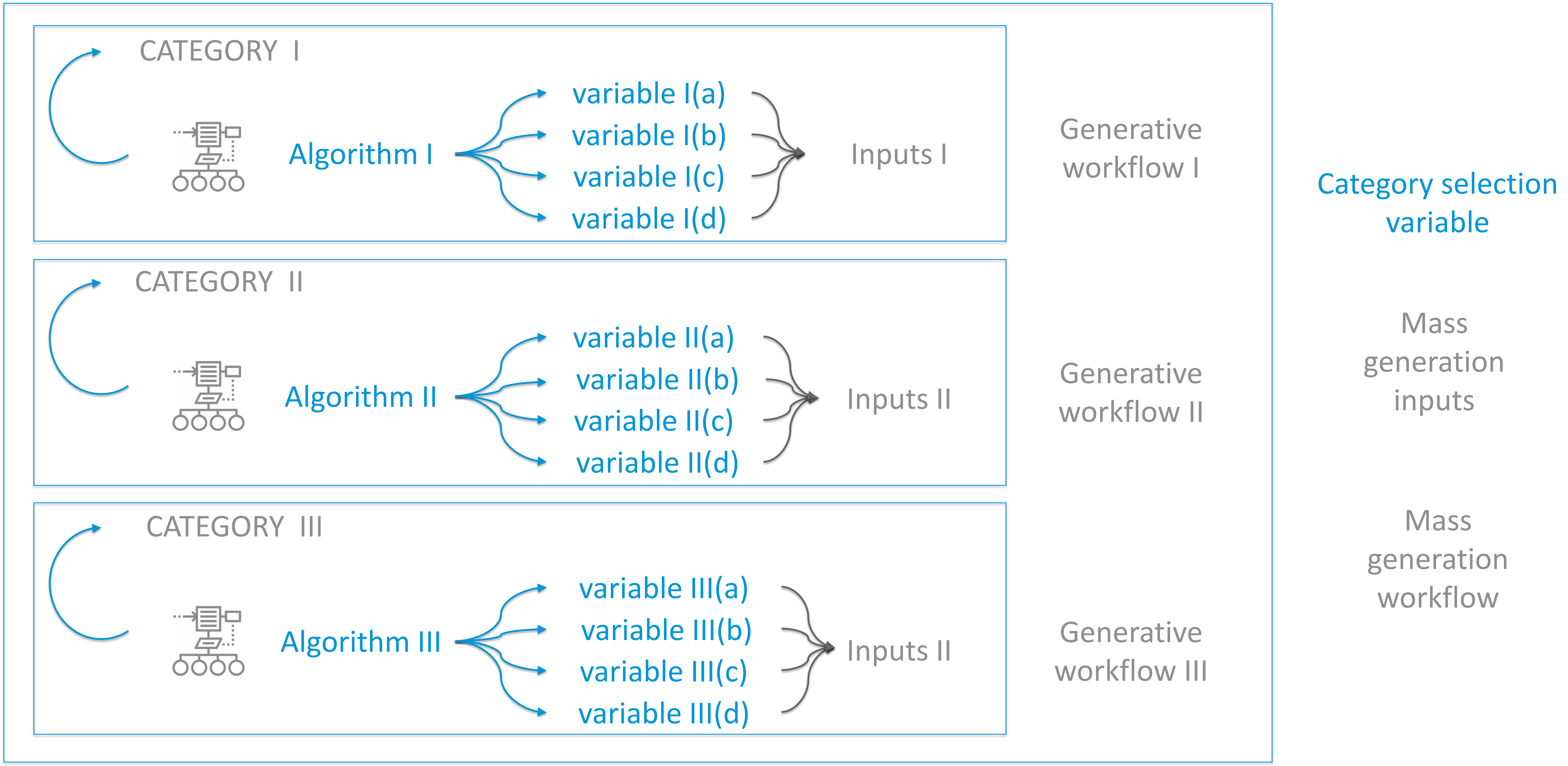
Avoid repetition / bias



1. Height
2. Width
3. Horizontal indent
4. Vertical indent

Algorithm

Synthetic building data : establishing variables

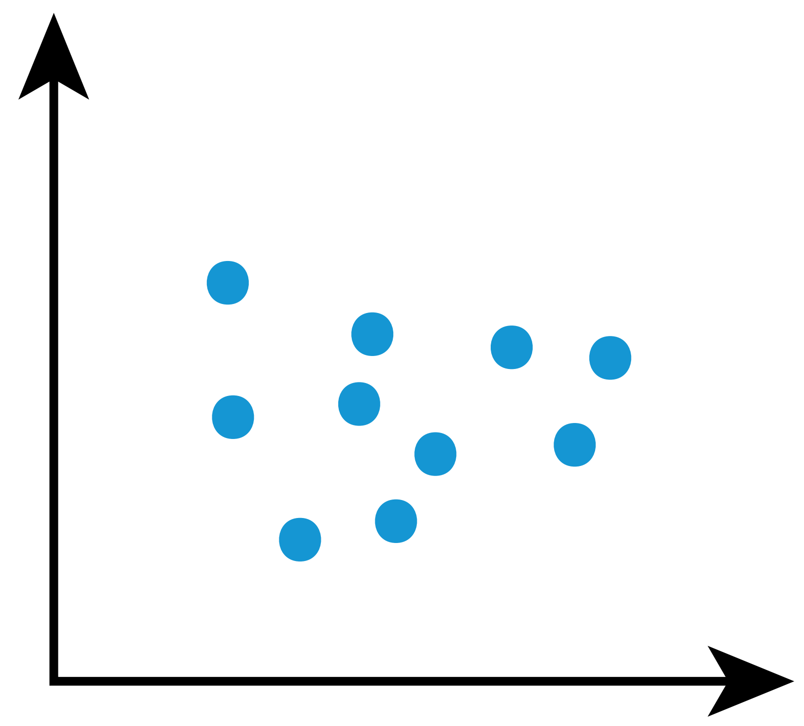


Synthetic building data : establishing objectives

Identify Dataset Objective → Identify Nature of Dataset → Choose Appropriate Solver → Generate Dataset

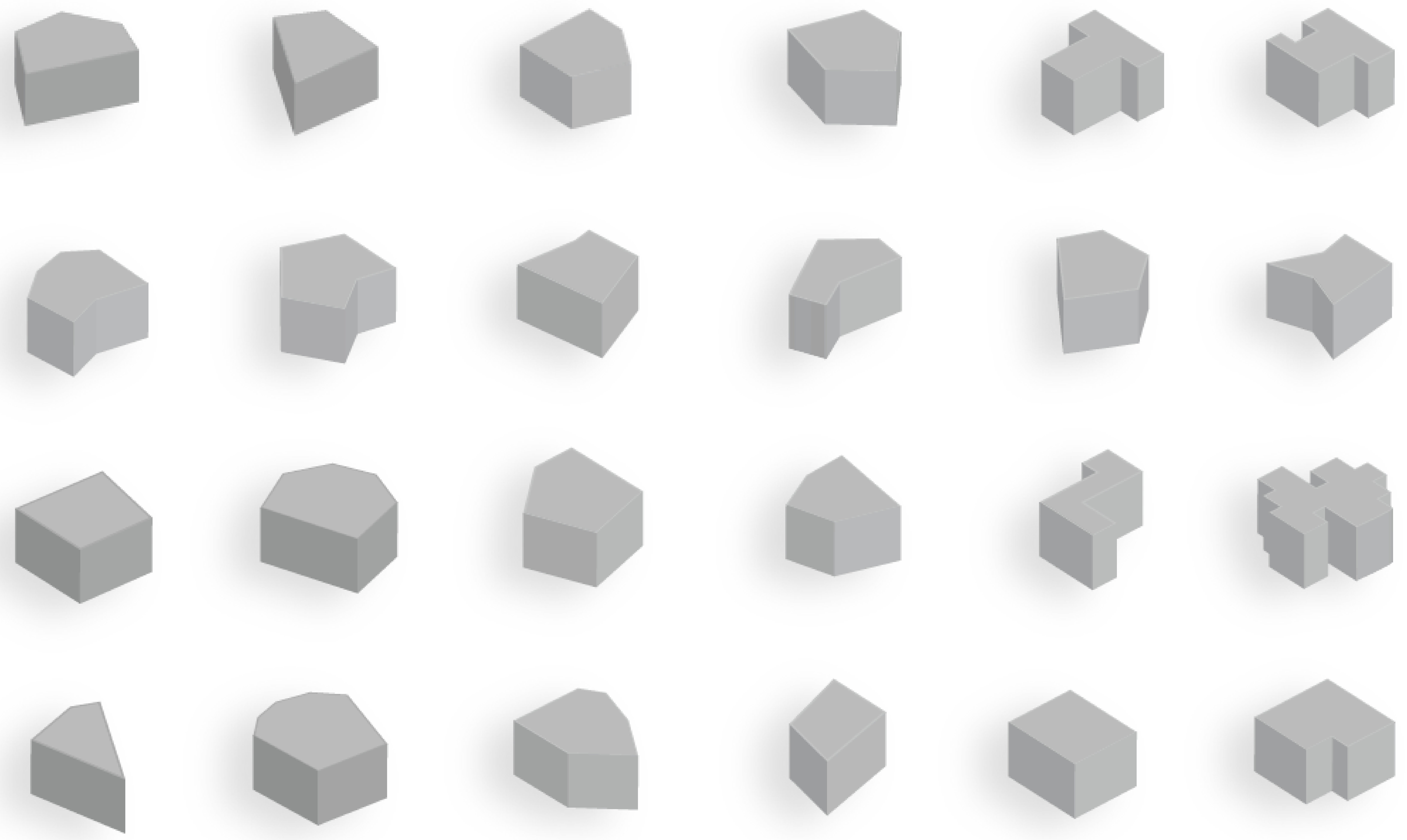
Synthetic building data : establishing objectives

RANDOMIZE SOLVER



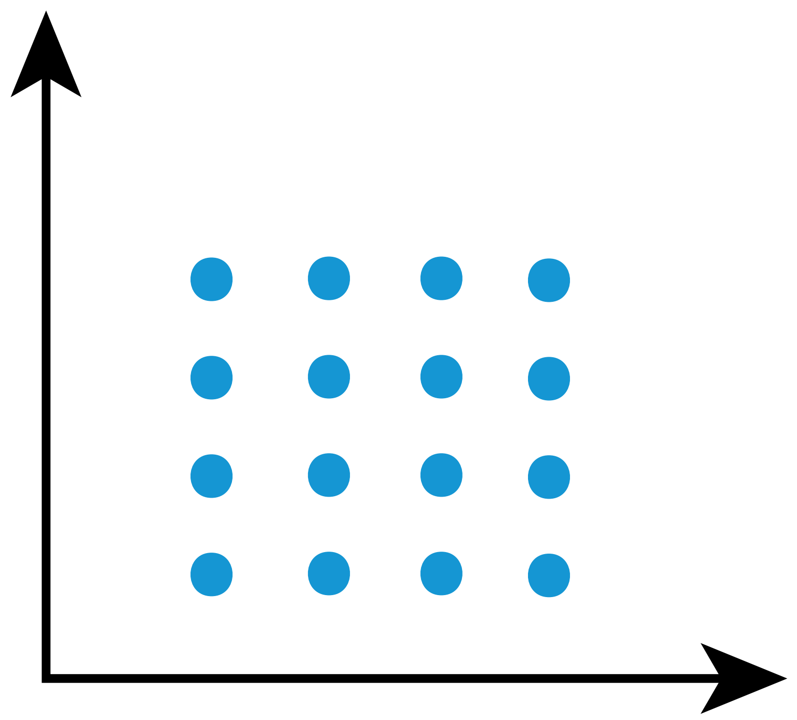
Sample space

Diverse datasets



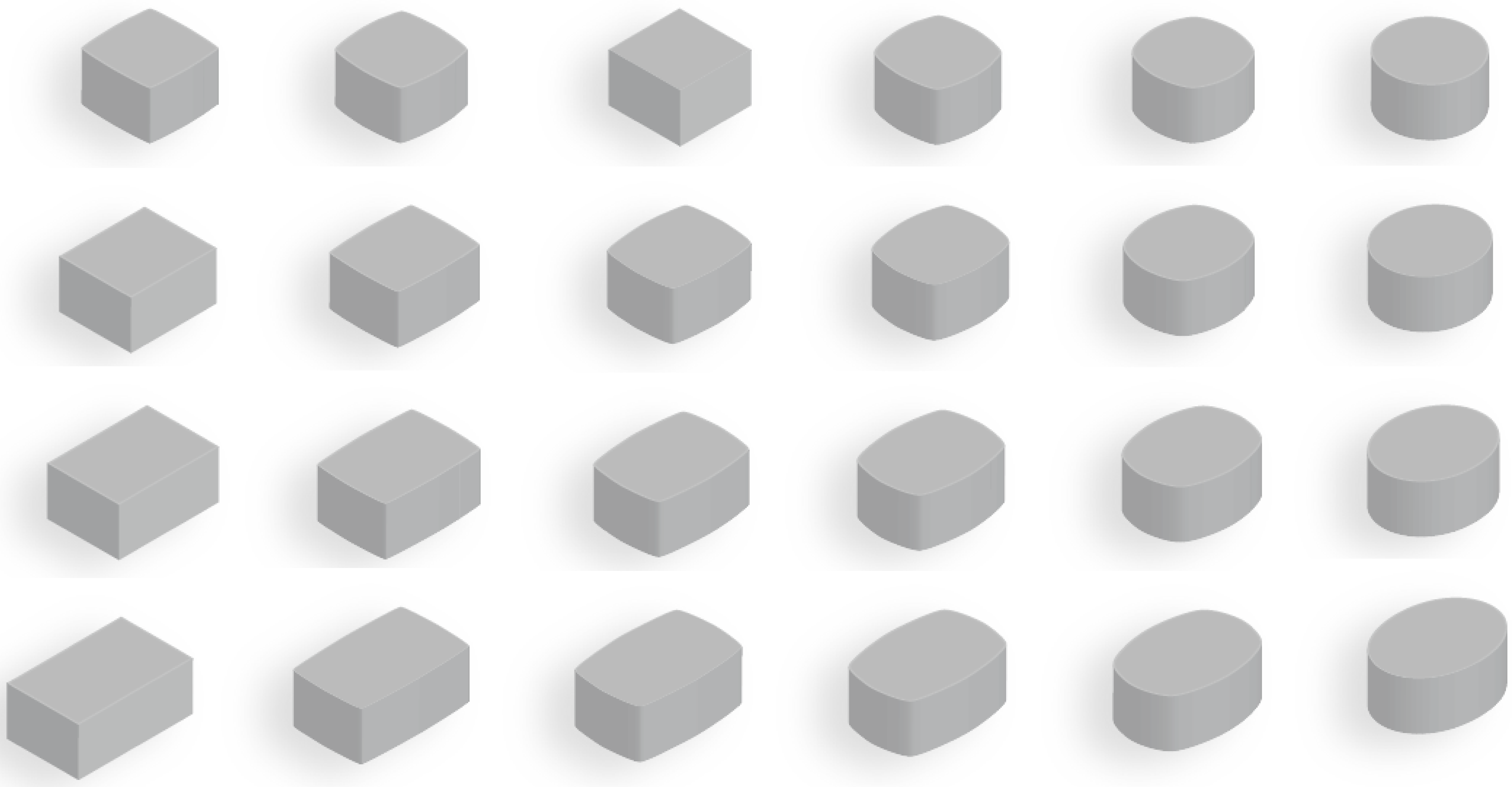
Synthetic building data : establishing objectives

CROSS-PRODUCT SOLVER



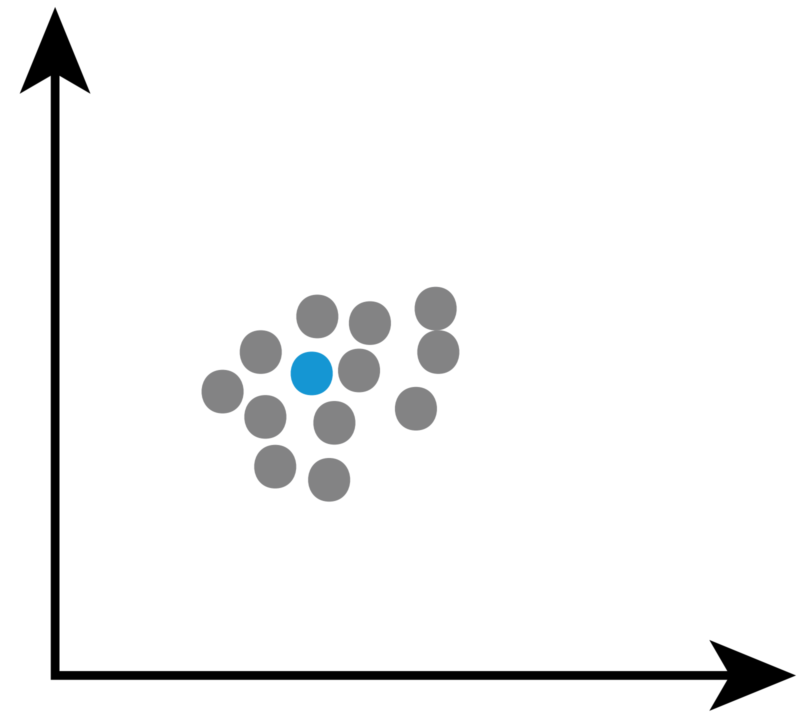
Sample space

Feature search/
Combinatorial analysis
dataset



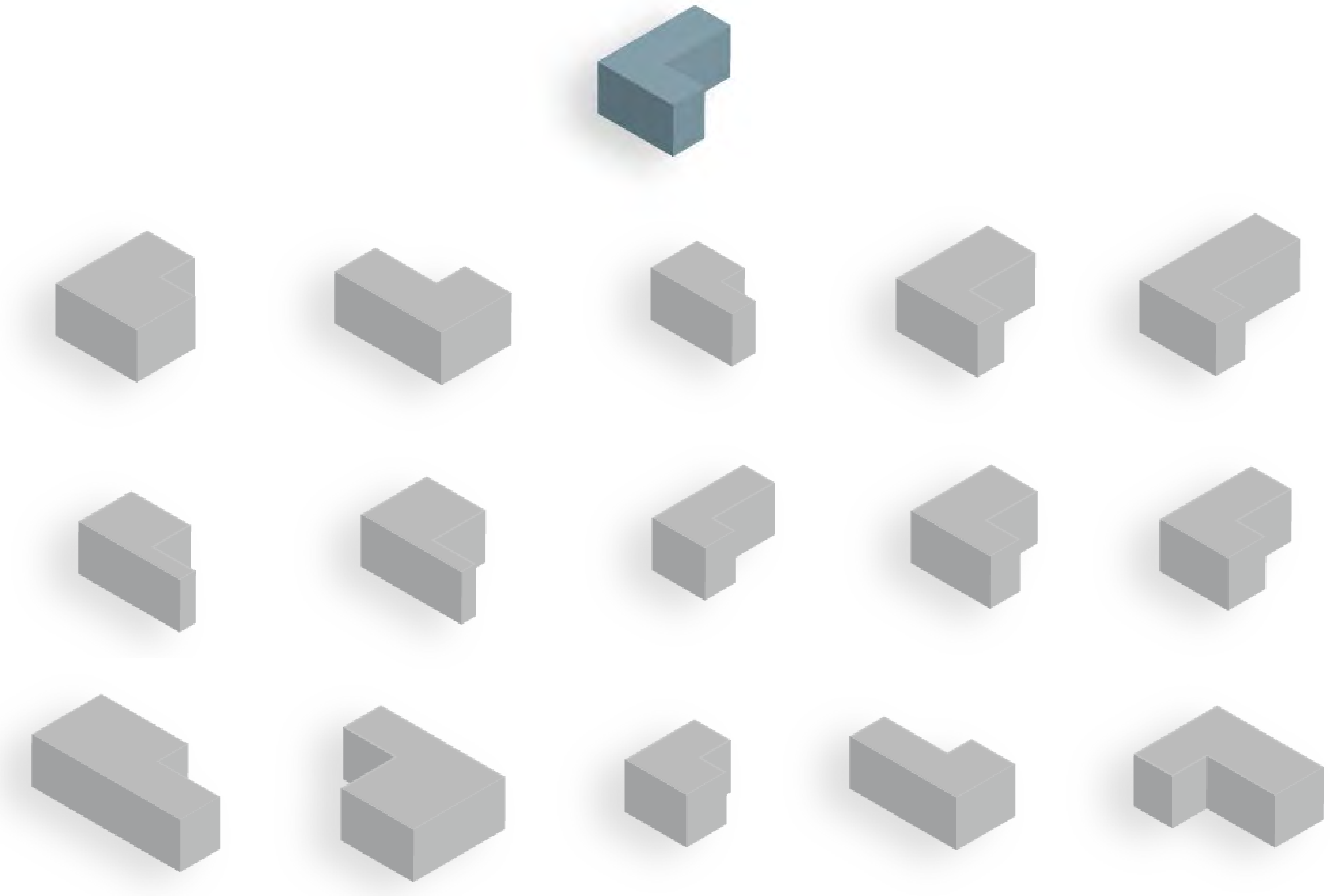
Synthetic building data : establishing objectives

LIKE – THIS SOLVER

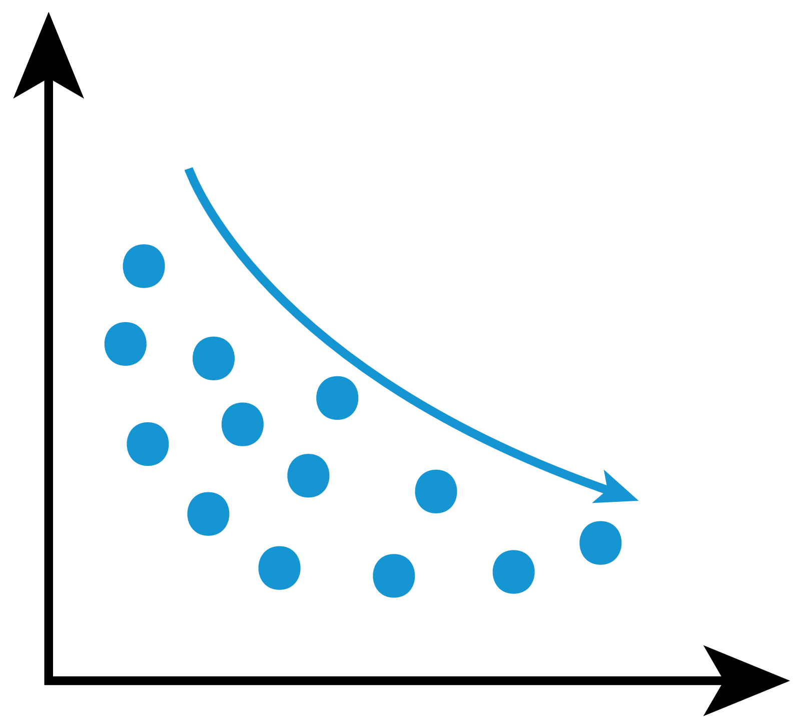


Sample space

Form finding / Sensitivity analysis dataset



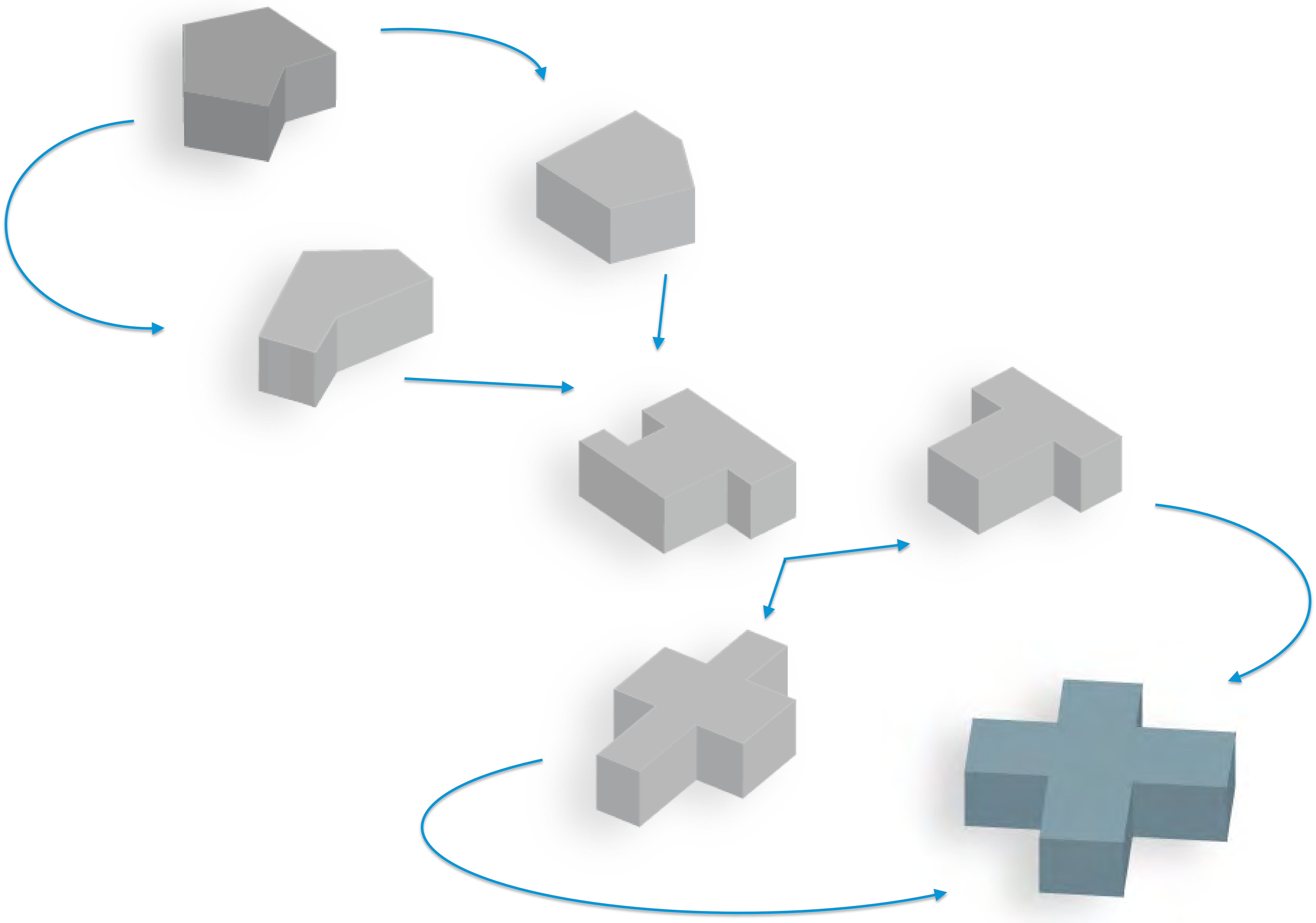
Synthetic building data : establishing objectives



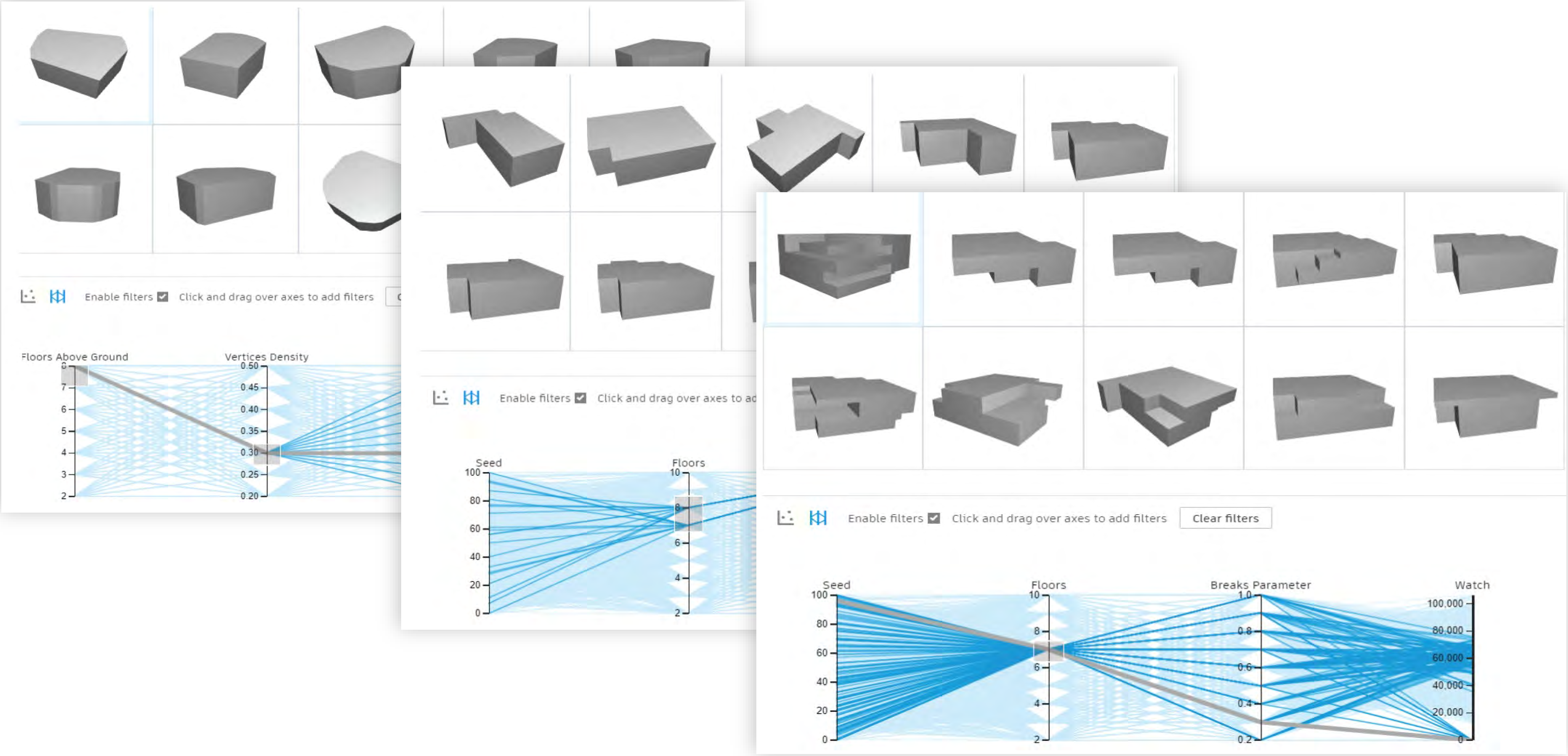
Sample space

Design Optimization / Performance oriented form-finding dataset

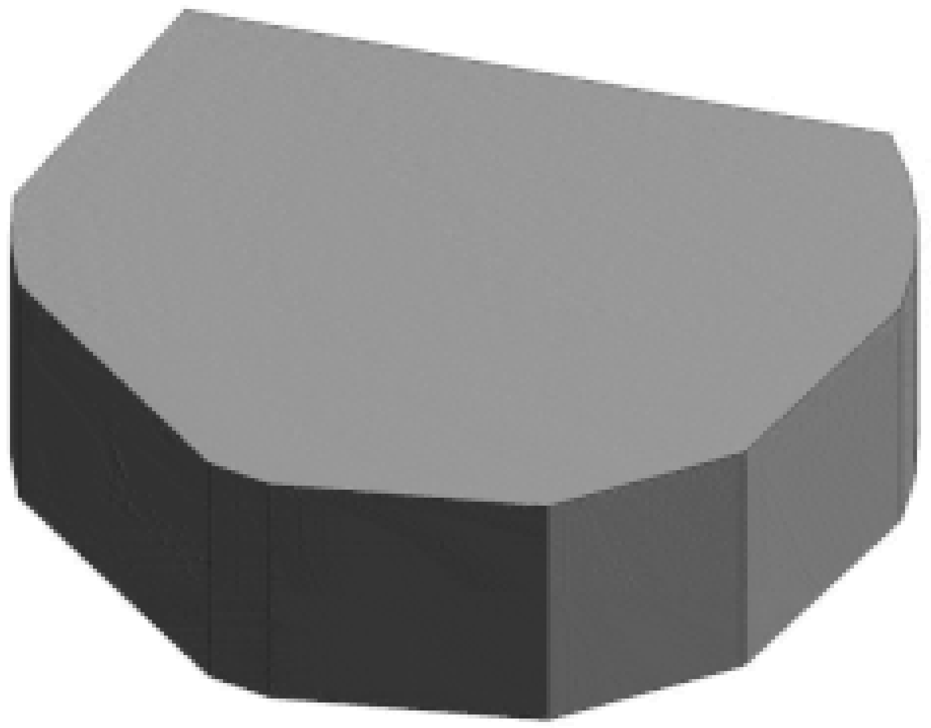
OPTIMIZE SOLVER



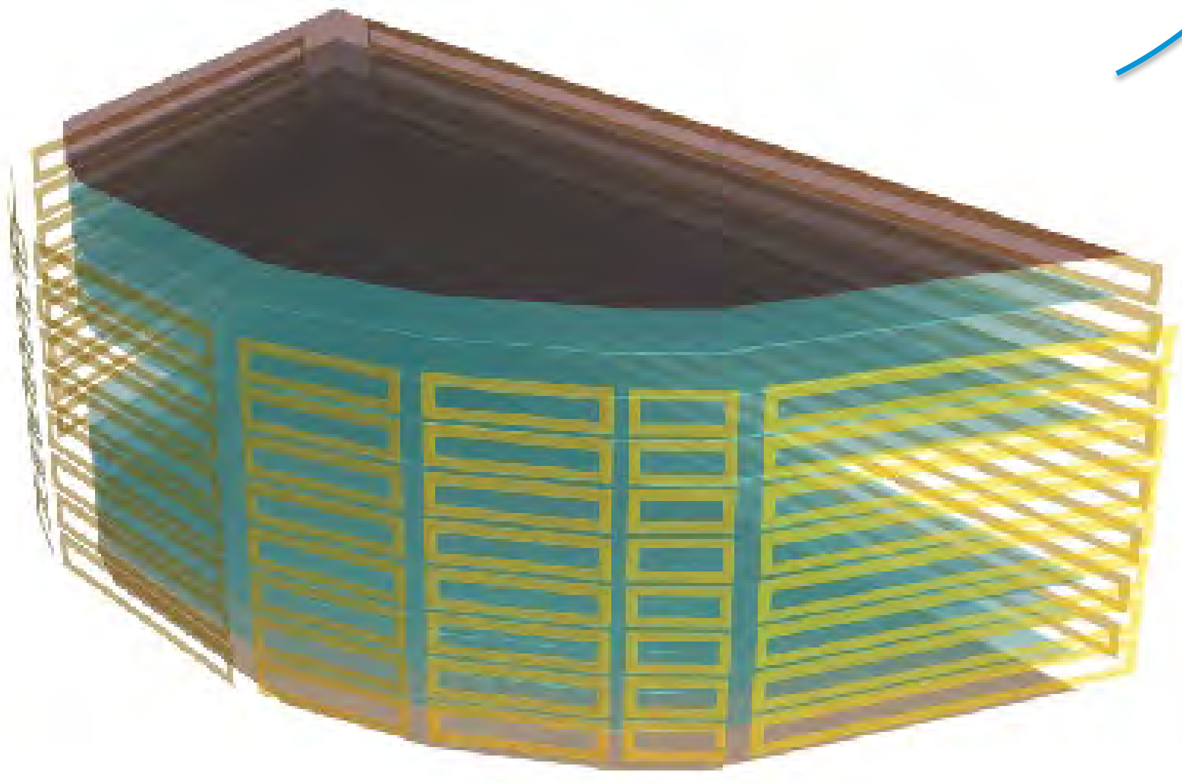
Synthetic building data : generating the dataset



Synthetic analytical data : ground truth labels



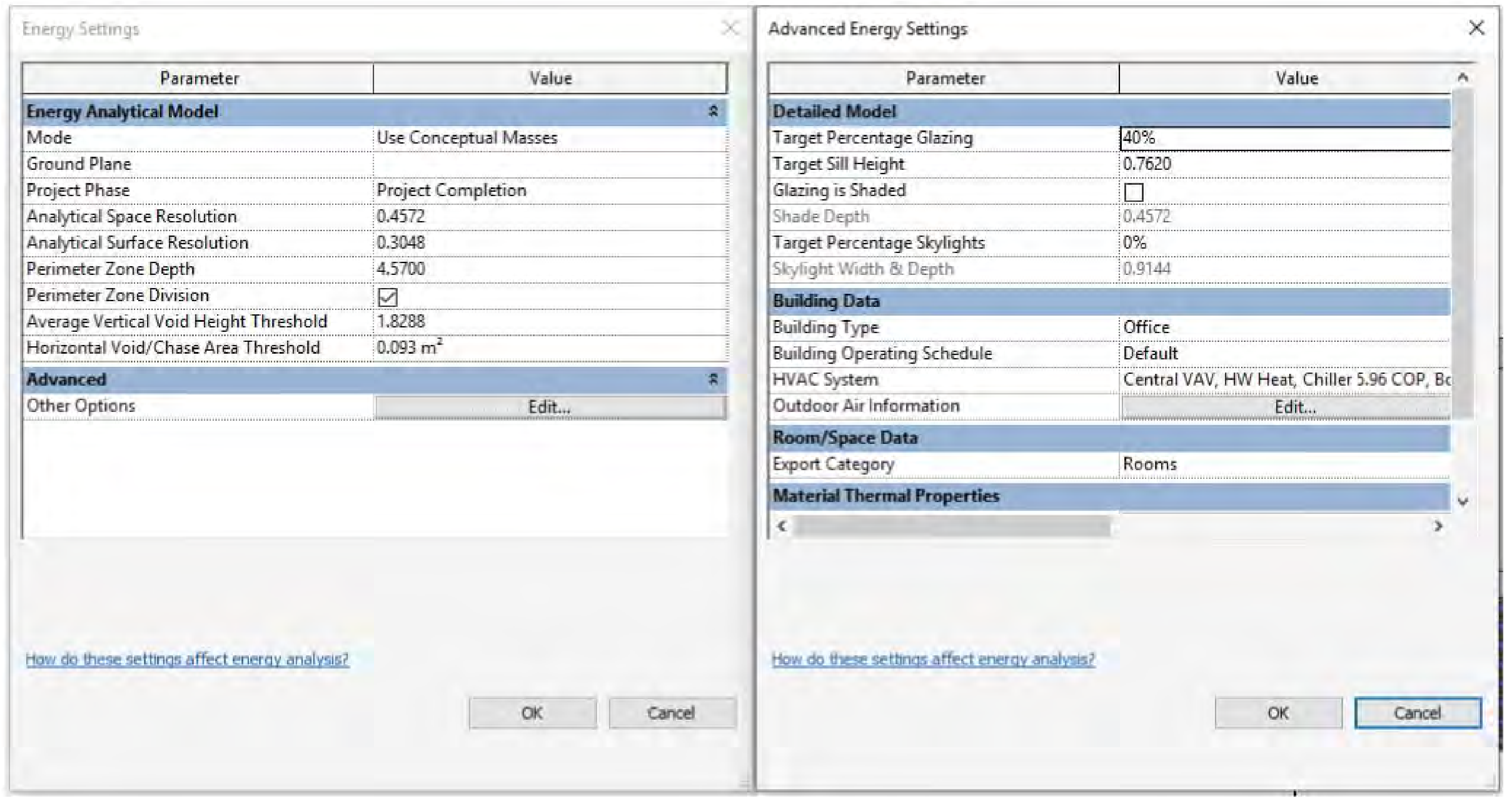
Generated solids



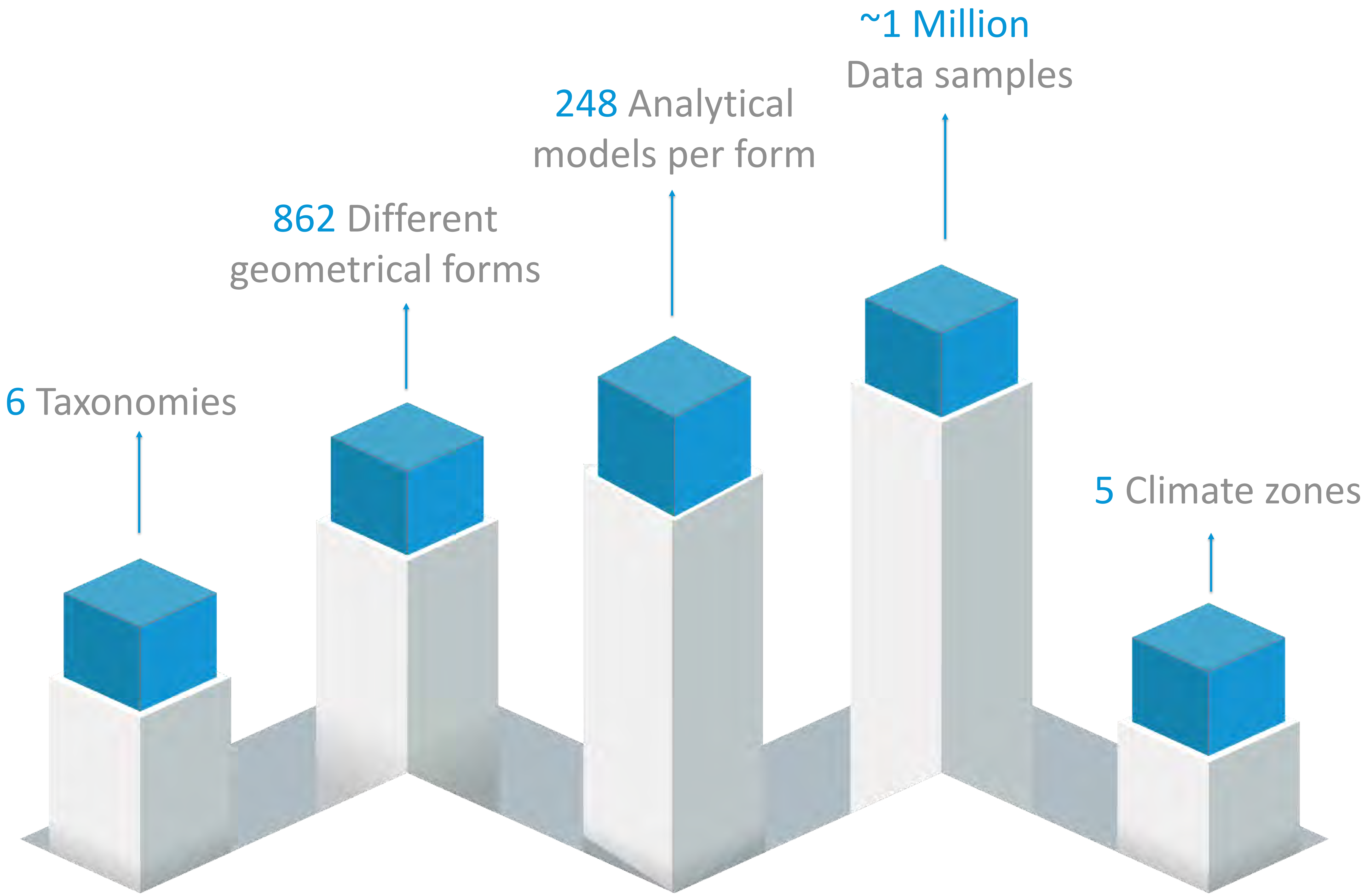
Analytical model
(GbXML)

Ground truth labels
(Simulation)

Revit energy settings



Synthetic data: a glance



Machine learning for performance prediction

Describing the problem

Describing the problem

Define the inputs and outputs

BUILDING

GEOMETRIC FORM
LOCATION
FACADES
CONSTRUCTIONS
OPERATION

→ **SIMULATION** → EUI

RESULT

physics based

number

Describing the problem

FEATURE VECTOR

X_0
X_1
X_2
…
X_n

x: number(s)

→ **SURROGATE MODEL** →

y = f(x)

RESULT

EUI

y: number

Representing the data

BUILDING

GEOMETRIC FORM
LOCATION
FACADES
CONSTRUCTIONS
OPERATION

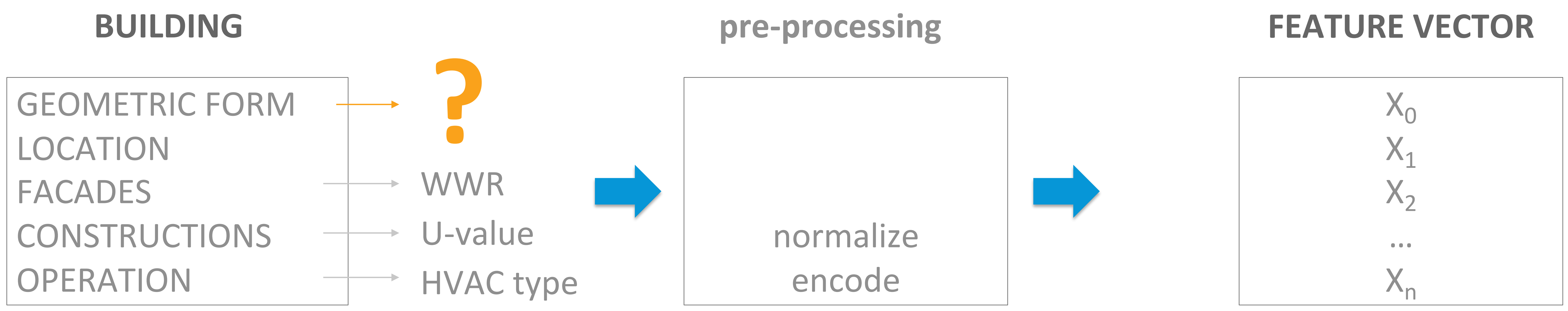
? →

FEATURE VECTOR

X_0
X_1
X_2
…
X_n

Representing the data

From buildings to feature vectors



Representing the data

Handling the Geometries

- **High Level Representations**
  - Parameters of the generative model
  - Post process geometries to extract parameters

selecting and
engineering features

- **Generic Geometric Format**
  - ~~Meshes ?~~
  - Images
  - Voxels
  - Point clouds

learning features

Model types

- Linear Regression
- Gaussian Process
- Random Forest
- Deep Learning
- …

Training a model

TRAINING SET

Fit model to data

VALIDATION SET

Model selection, tuning

TEST SET

Evaluate final model

How to start?

SELECT MODELS, DATA REPRESENTATIONS

Use domain knowledge

Use precedents

BUILD DATASET

Identify sources

Combine data

Cleanup data

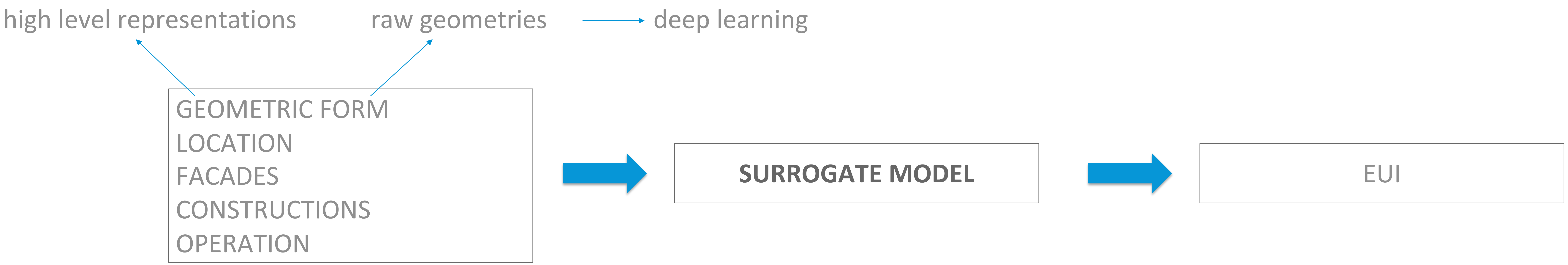
TEST AND REPEAT

Simplify problem

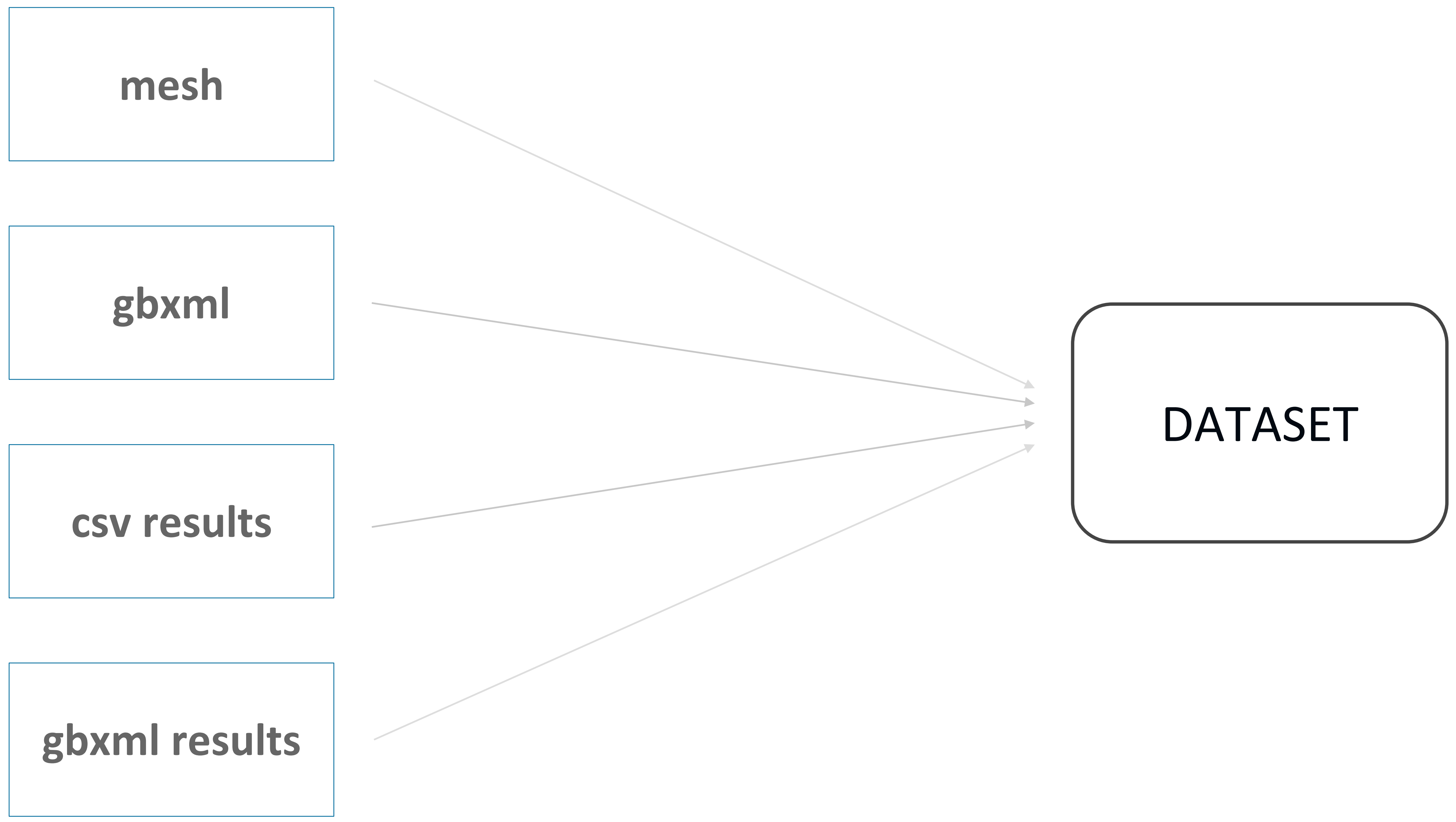
Use subset of data

Add complexity

Annual EUI prediction



Data sources



Annual EUI prediction

Annual EUI prediction

SIMPLE GEOMETRIES → **SURROGATE MODEL** → EUI

Results inform the:

model type

geometry representations

dataset diversity

Annual EUI prediction

COMPLEX GEOMETRIES → **SURROGATE MODEL** → EUI

Annual EUI prediction

~862 geometric forms

X

242 parameter variations

X

5 climate zones

1.91 % error

Annual EUI prediction

Get annual EUI result in milliseconds

Dynamo
BPA_MachinePrototype.dyn*
Design Parameters
Height Change
Length Change
Width Change
Python Script
Insight Parameters
Zone
WWR
Shading %
Glazing type
Wall Construction
Roof Construction
Infiltration
Lighting Efficiency
PLug-load Efficiency
Daylight + Occupancy Controls
HVAC
Operating Schedule
Python Script || Insight
Parameters
PredictEUI.GetEUI
Geometry
Zone
WWRNorth
WWRSouth
WWREast
WWRWest
WindowShadingNorth
WindowShadingSouth
WindowShadingEast
WindowShadingWest
WindowGlazingNorth
WindowGlazingSouth
WindowGlazingEast
WindowGlazingWest
WallConstruction
RoofConstruction
Infiltration
LightingEfficiency
PlugLoadEfficiency
LightingAndOccupancyControls
HVAC
OperationSchedule
double
EUI
Library
Dictionary
Display
Geometry
ImportExport
Input
List
Math
Revit
Script
String
EnergyPrediction
Automatic
Properties
3D View
View Scale
Custom
Scale Value 1:
96
Detail Level
Medium
Parts Visibility
Show Original
Discipline
Architectural
Project Browser - Project1
Views (all)
Floor Plans
level 0
3D Views
01 - Existing
02 - Demo
A10 - Substructure
Analytical Spaces
B10 - Superstructure
B20 - Exterior Enclosure
C10 - Interior Constructio
C20 - Interior Finish
E20 - Furnishings
Perspective 3D
System-Zones
{3D}
Elevations (Building Elevation
East
North
South
West
Main Model
Type here to search
11:21 PM
10/19/2020

To Sum Up

[1]

Discussed the importance and current obstacles of integrating energy prediction early in design.

[2]

Introduced the potential of ML methods as a solution, and the necessity of synthetic data.

[3]

Demonstrated an example of generating a diverse dataset of labeled synthetic data.

[4]

Introduced the steps for training and evaluating a ML model and demonstrated how achieved real-time EUI predictions.

Please put any questions you have in the class comment box!

Michael Floyd · michael.floyd@autodesk.com

Varvara Toulkeridou · varvara.toulkeridou@autodesk.com

Vishal Vaidhyanathan · vishal.vaidhyanathan@autodesk.com

Spyridon Ampanavos · sampanavos@gsd.harvard.edu