

# PhysX-3D: Physical-Grounded 3D Asset Generation

Ziang Cao<sup>1</sup>, Zhaoxi Chen<sup>1</sup>, Liang Pan<sup>2</sup>, Ziwei Liu<sup>1</sup>

<sup>1</sup>Nanyang Technological University

<sup>2</sup>Shanghai AI Lab



# Motivation

- Absence of Physical-Grounded 3D Datasets.

Most existing 3D datasets primarily provide geometry and texture information, but they rarely include **physical properties** such as **absolute scale**, **material**, **affordance**, **kinematics**, and **function description**, thereby constraining the development of generative models.

Dataset	# Objs	Part anno	Physical Dim	Material	Affordance	Kinematic	Description	Year
ShapeNet [3]	51K	✗	✗	✗	✗	✗	✗	2015
PartNet [17]	26K	✓	✗	✗	✗	✗	✗	2019
PartNet-Mobility [26]	2.7K	✓	✗	✗	✗	✓	✗	2020
GAPartNet [9]	1.1K	✓	✗	✗	✗	✓	✗	2022
ABO [6]	7.9K	✗	✓	Obj-level	✗	✗	Obj-level	2022
OmniObject3D [25]	6K	✗	✗	✗	✗	✗	✗	2023
Objaverse [8]	818K	✗	✗	✗	✗	✗	✗	2023
PhysXNet (ours)	26K	✓	✓	Part-level	✓	✓	Part-level	2025
PhysXNet-XL (ours)	6M	✓	✓	Part-level	✓	✓	Part-level	2025

# Motivation

- Limitations in current 3D generative methods.

Existing 3D generation primarily emphasizes geometries and textures while **neglecting physical-grounded modeling**, hampering their real-world application in physical domains like simulation and embodied AI.

Image Prompts



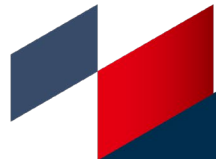
Geometry and appearance



Image Prompts

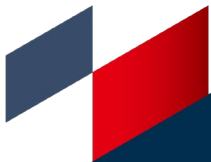


Geometry and appearance



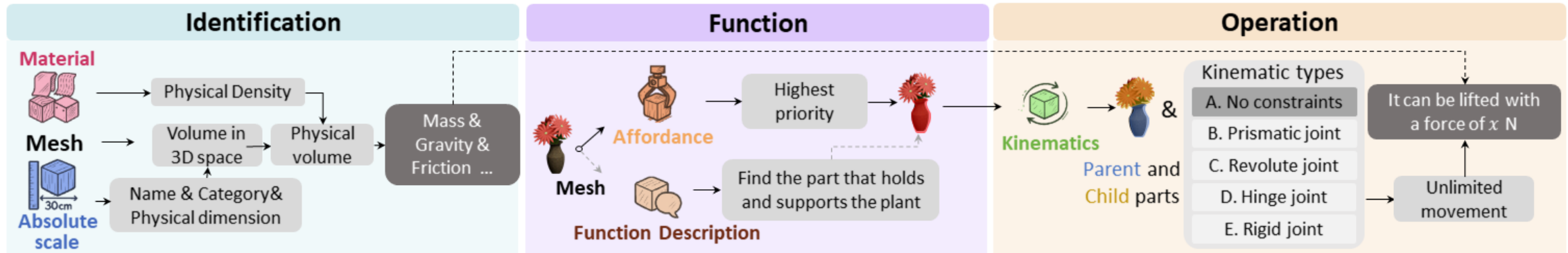
# Contribution

- We pioneer the **first end-to-end paradigm** for physical-grounded 3D asset generation, advancing the research frontier in physical-grounded content creation and unlocking new possibilities for downstream applications in simulation.
- We build the **first physical-grounded 3D dataset, PhysXNet**, and propose a **human-in-the loop annotation** pipeline to convert existing geometry-focused datasets into fine-grained physics-annotated 3D datasets efficiently and robustly. In addition, we present an extended version, **PhysXNet-XL**, which includes over 6 million annotated 3D objects generated through procedural methods.
- We design a dual-branch feed-forward framework, **PhysXGen**. It can model the latent interdependencies between structural and physical features to achieve plausible physical predictions while maintaining the native geometry quality.



# Definition of physical properties

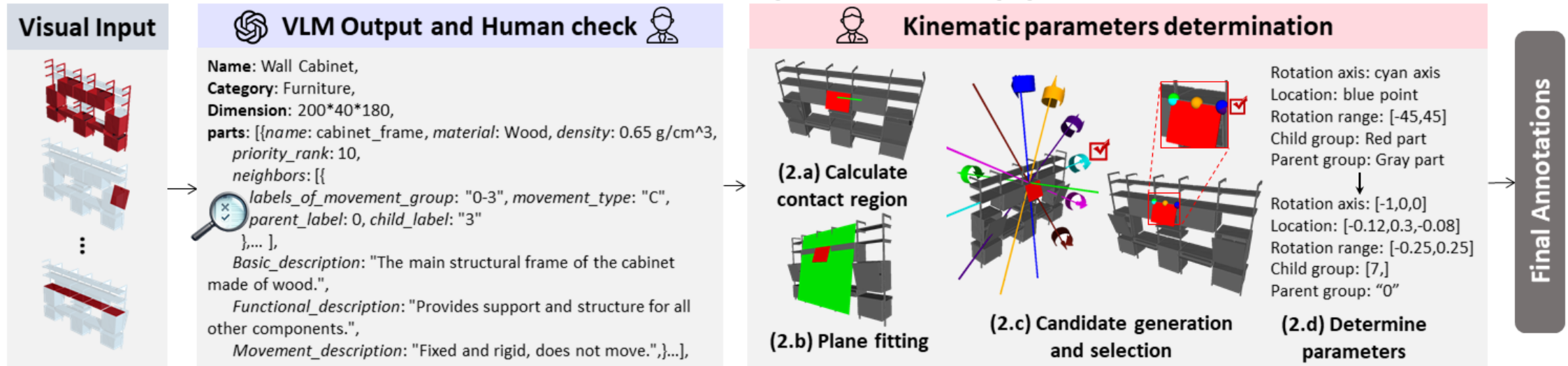
## Definition of physical and semantic properties



- **Identification** - determining the basic nature of the object  
absolute scaling and material (material name, Young's modulus, Poisson's ratio, and density)
- **Function** - understanding its potential applications  
functional affordance analysis and function descriptions (basic, functional, and kinematic descriptions).
- **Operation** - detailed usage methodologies  
kinematic parameter

# HITL annotation pipeline

## Human-in-the-loop annotation pipeline



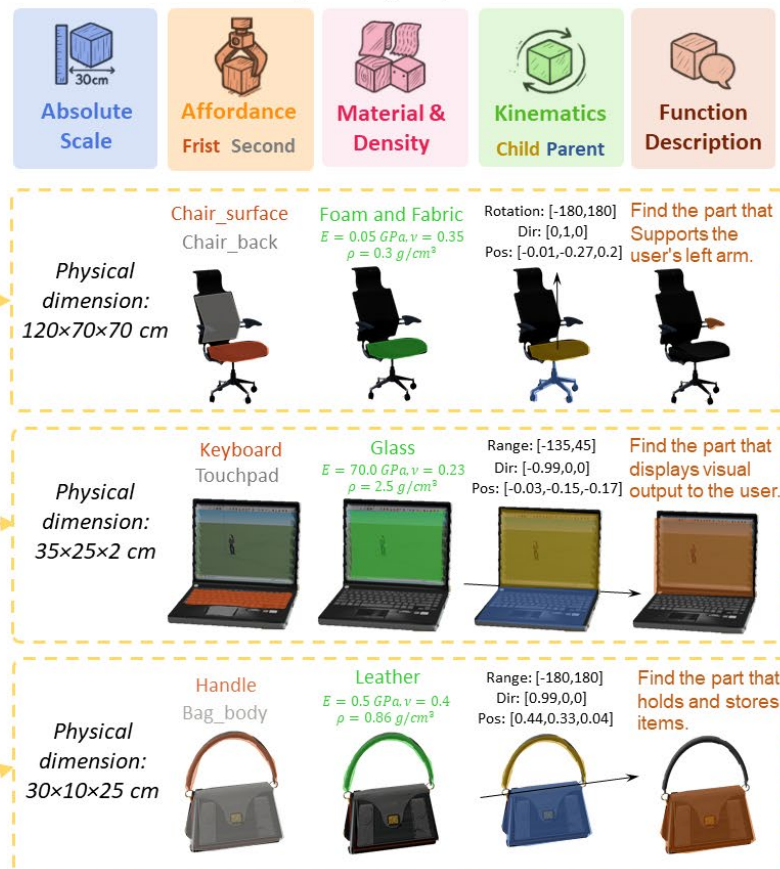
- **Preliminary Data Acquisition:**
  - Utilize GPT-4o to obtain the basic information
  - Human Check
- **Kinematic Parameter Determination**
  - Calculate contact region
  - Plane fitting
  - Candidate generation and selection
  - Determine kinematic parameters by human



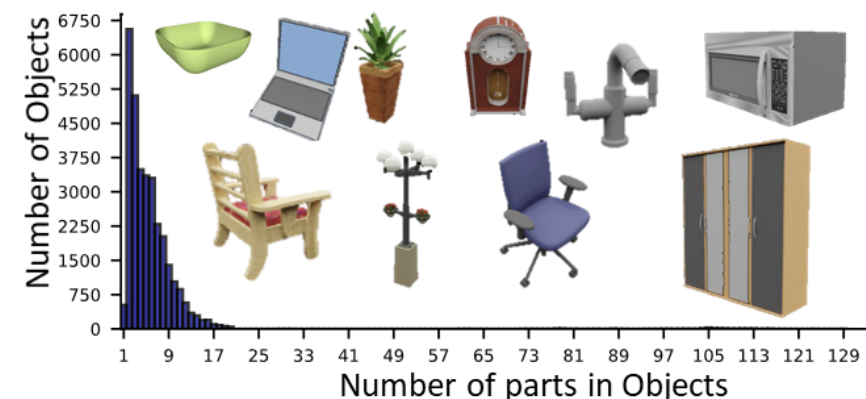
# PhysXNet&PhysXNet-XL



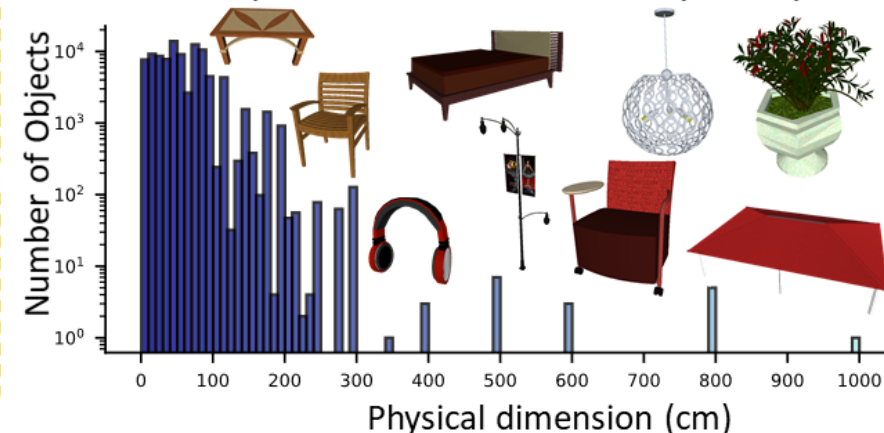
## Physical properties



## Part Number Occurrence Frequency



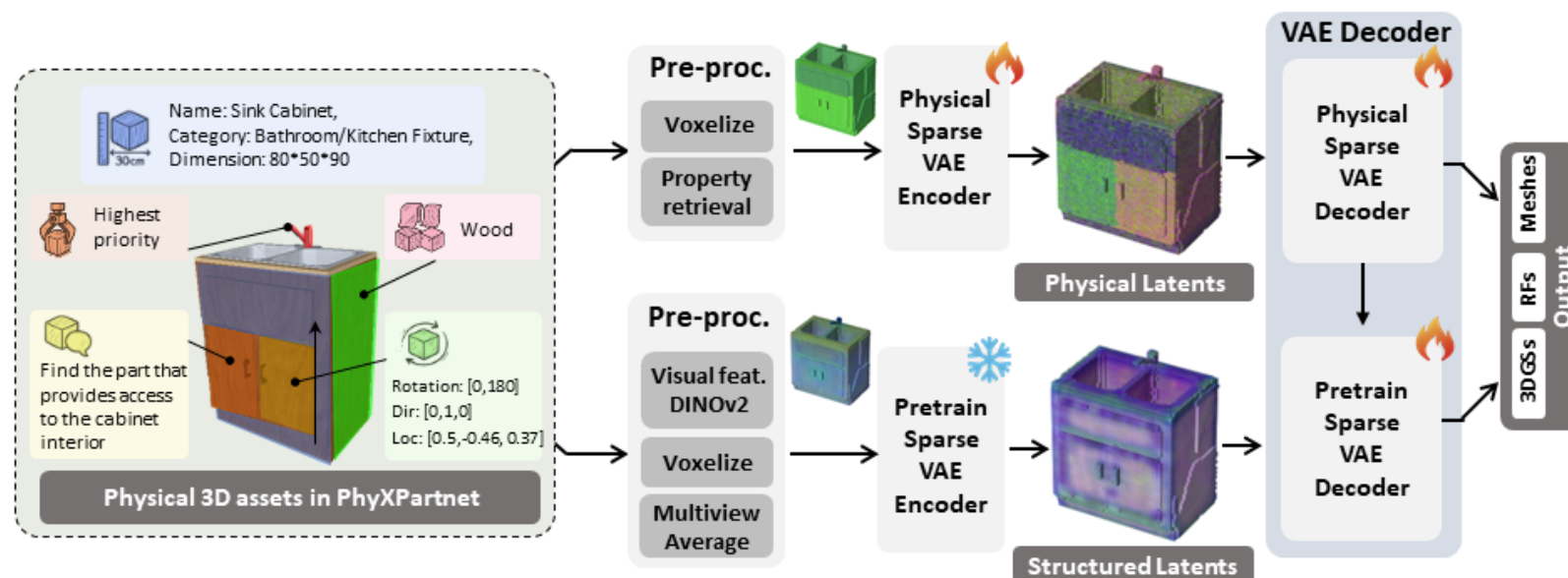
## Physical Dimension Frequency



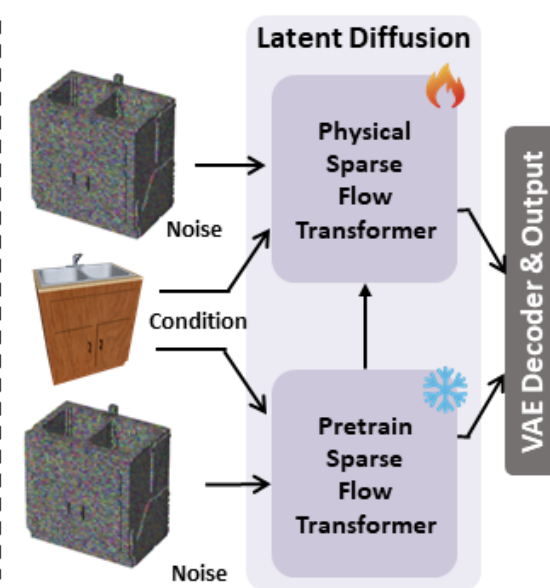
We propose **PhysXNet&PhysXNet-XL** – the first comprehensive physical 3D dataset containing over **26K&6M** richly annotated 3D objects

# PhysXGen

## Physical 3D Assets VAE Encoding & Decoding



## Physical Latent Generation



PhysXGen features a two-stage architecture comprising:

- physical 3D VAE framework for latent space learning
- physics-aware diffusion process for structured latent generation.



# Experimental results

Image Prompts



Geometry and appearance



Physical properties

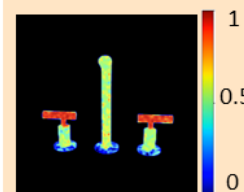
Absolute Scale

Physical  
dimension:  
 $27.51 \times 19.8 \times 6.76$   
cm

Material

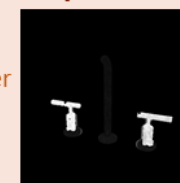
$\rho \approx 8.2 \text{ g/cm}^3$

Affordance



Function Description

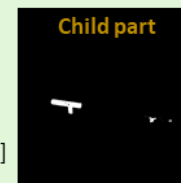
Find the part that  
Used to turn water  
on/off or adjust  
temperature.



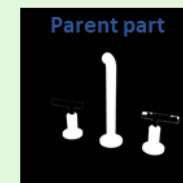
Kinematics

Kinematic type: rotation  
Range:  $[-92.3, 87]$   
Dir:  $[0.18, 0.736, 0.02]$   
Pos:  $[-0.56, -0.04, -0.073]$

Child part



Parent part



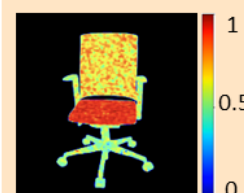
Absolute Scale

Physical  
dimension:  
 $98.92 \times 69.2 \times 64.3$   
cm

Material

$\rho \approx 8.0 \text{ g/cm}^3$

Affordance



Function Description

Find the mesh  
fabric backrest  
surface of the chair



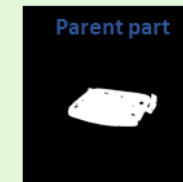
Kinematics

Kinematic type: rotation  
Range:  $[-190.8, 143.1]$   
Dir:  $[0.02, 0.864, -0.03]$   
Pos:  $[0.032, -0.11, 0.11]$

Child part



Parent part

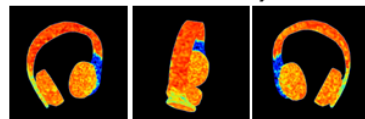


# Experimental results

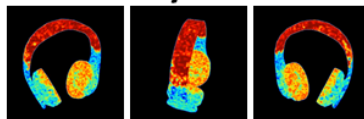
## Prompts



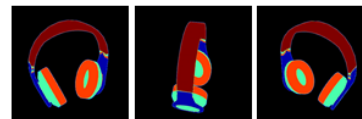
TRELLIS+PhysPre



PhysXGen



Ground Truth



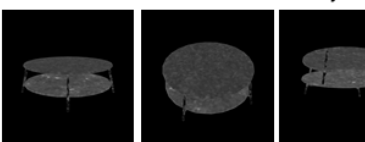
## Property



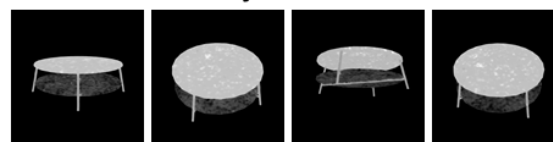
Affordance



TRELLIS+PhysPre

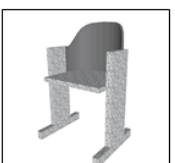


PhysXGen



Function  
Description

Find the part that serves as the main surface for placing items



TRELLIS+PhysPre

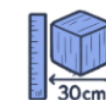
Physical Dimension:  
72.3×71.5×97.1 cm

PhysXGen

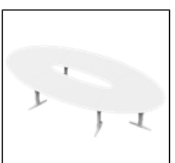
Physical Dimension:  
61.2×62.4×88.9 cm

Ground Truth

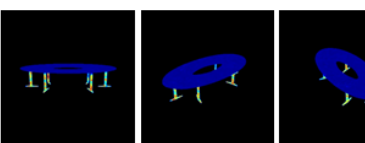
Physical Dimension:  
60×60×90 cm



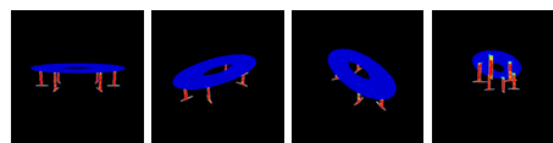
Absolute  
Scale



TRELLIS+PhysPre



PhysXGen



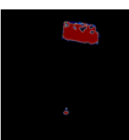
Material &  
Density

0.583 g/cm<sup>3</sup>

3.448 g/cm<sup>3</sup>

7.48 g/cm<sup>3</sup>

TRELLIS+PhysPre



Child part Parent part

Type: E

Range: [-174.6, -169.2]

Dir: [-0.42, -0.53, 0.26]

Pos: [0.32, 0.16, 0.52]

PhysXGen



Child part Parent part

Type: C

Range: [-159.8, 160.8]

Dir: [0.13, 0.05, 0.86]

Pos: [0.06, 0.01, 0.83]



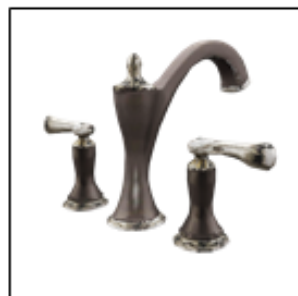
Kinematics

# Experimental results

Image prompts

Trellis+PartField+GPT-4o

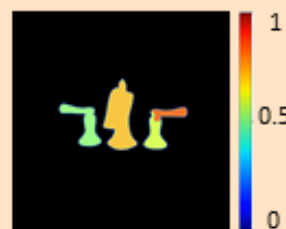
PhysXGen



Absolute Scale

Physical  
dimension:  
20×15×12 cm

Affordance



Material

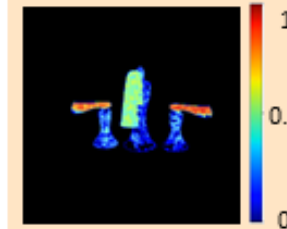
$\rho \approx 7.8 \text{ g/cm}^3$



Absolute Scale

Physical  
dimension:  
24.31×18.19×14.7  
cm

Affordance

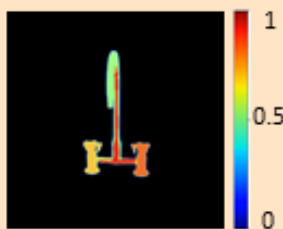


Material

$\rho \approx 7.8 \text{ g/cm}^3$

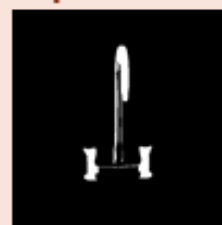


Affordance

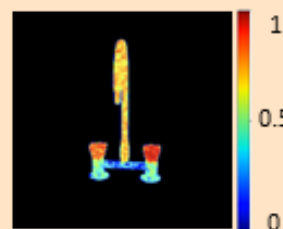


Function Description

Find the part that  
delivers water  
from the valve  
system to the sink

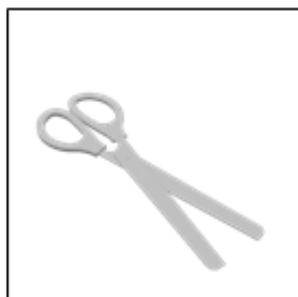
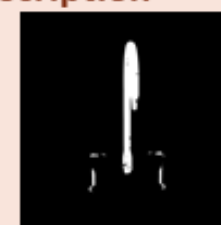


Affordance



Function Description

Find the part that  
delivers water  
from the valve  
system to the sink



Kinematics

Kinematic type: rotation  
Range: [-180,180]  
Dir: [1,0,0]  
Pos: [0,0,0]

Child part



Parent part



Kinematics

Kinematic type: rotation  
Range: [-87.3,10.2]  
Dir: [-0.01,0.941,0.05]  
Pos: [-0.011,0.013,0.001]

Child part



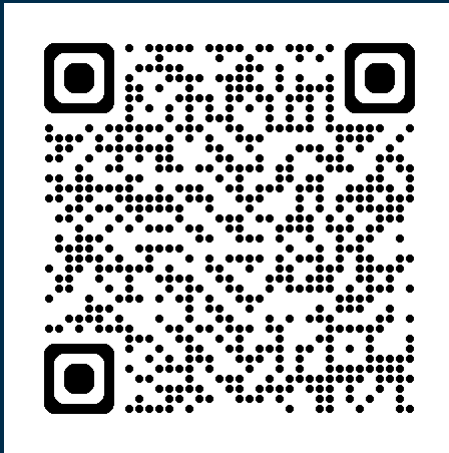
Parent part



# PhysX-3D: Physical-Grounded 3D Asset Generation

Ziang Cao<sup>1</sup>, Zhaoxi Chen<sup>1</sup>, Liang Pan<sup>2</sup>, Ziwei Liu<sup>1</sup>

<sup>1</sup>Nanyang Technological University, <sup>2</sup>Shanghai AI Lab



Home page:

<https://physx-3d.github.io/>

**Physical properties**

Absolute Scale	Affordance	Material & Density	Kinematics	Function Description
	Frist Second		Child Parent	

Chair_surface Chair_back	Foam and Fabric $E = 0.05 \text{ GPa}, \nu = 0.35$ $\rho = 0.3 \text{ g/cm}^3$	Rotation: [-180,180] Dir: [0,1,0] Pos: [-0.01,-0.27,0.2]	Find the part that Supports the user's left arm.
Physical dimension: 120×70×70 cm			

Keyboard Touchpad	Glass $E = 70.0 \text{ GPa}, \nu = 0.23$ $\rho = 2.5 \text{ g/cm}^3$	Range: [-135,45] Dir: [-0.99,0,0] Pos: [-0.03,-0.15,-0.17]	Find the part that displays visual output to the user.
Physical dimension: 35×25×2 cm			

Handle Bag_body	Leather $E = 0.5 \text{ GPa}, \nu = 0.4$ $\rho = 0.86 \text{ g/cm}^3$	Range: [-180,180] Dir: [0.99,0,0] Pos: [0.44,0.33,0.04]	Find the part that holds and stores items.
Physical dimension: 30×10×25 cm			