

# A2Z-10M+: Geometric Deep Learning with A-to-Z BRep Annotations for AI-Assisted CAD Modeling and Reverse Engineering

## Supplementary Material

### Contents

<b>A Multi-scale SPH for Neighbor Labeling</b>	<b>1</b>
A.1 Coedge Length–Aware Neighbors. . . . .	1
A.2 SPH kernel with BRep Surface features. .	1
<b>B Synthesized Sketch and Realism Analysis</b>	<b>2</b>
B.1 Generation Pipeline . . . . .	2
B.2 Quantitative Realism Evaluation . . . . .	2
<b>C Synthesized Scan and Realism Analysis</b>	<b>2</b>
C.1 Artec3D-Calibrated Scan Synthesis . . .	2
C.2 Stage-Wise Realism Evaluation . . . . .	3
<b>D Prompt for CAD Description and Tags</b>	<b>3</b>
D.1 Description . . . . .	4
D.2 Applying Filtering and Constraints . . .	4
D.3 Output Format . . . . .	4
<b>E Hierarchical Tags for CAD Retrieval</b>	<b>4</b>
E.1 Semantic Hierarchy over 1M A2Z Models	5
E.2 Heterogeneous Graph Representation . .	6
E.3 Graph Mining and Retrieval Mechanism .	6
<b>F GPT-5, Gemini, and Human Evaluation.</b>	<b>6</b>
F.1. Chunk-wise Evaluation of Model . . . . .	6
<b>G More Examples of Caption and Tags</b>	<b>7</b>
<b>H More Examples of Electronic Enclosures</b>	<b>7</b>
<b>I. Examples of Annotations and Predictions</b>	<b>7</b>

We structure the supplementary material according to the content outlined above.

### A. Multi-scale SPH for Neighbor Labeling

In A2Z, we have generated 3D meshes that closely resemble real-world scanning outcomes. This process is carried out by uniquely combining multiple computational geometry algorithms, guided by BRep geometry and topology features. Many uncommon problems arise when too many small faces are adjacent to one another. These corner cases create confusion about how to assign the nearest neighbor. Let  $\mathcal{P} = \{\mathbf{p}_i\}_{i=1}^N$  be the scan, and let  $\mathbf{x} \in \mathcal{S}_E \cup \mathcal{S}_F$  be a sampled BRep point with local frame  $R(\mathbf{x}) = [\mathbf{t}, \mathbf{u}, \mathbf{n}]$  (tangent, in-plane, normal).

#### A.1. Coedge Length{Aware Neighbors.

For a boundary sample  $\mathbf{x} \in \mathcal{S}_E$  with parent edge length  $L_{e(\mathbf{x})}$ , we first define a base scale  $h_E(\mathbf{x}) = \alpha_E L_{e(\mathbf{x})}$ ,  $\mathcal{R}_E(\mathbf{x}) = \{r_k(\mathbf{x}) = \gamma_k h_E(\mathbf{x})\}_{k=1}^K$ , and then for a BRep face sample point  $\mathbf{x} \in \mathcal{S}_F$  with grid pitch  $\Delta_f$ , the same base scale for the face is defined as:

$$h_F(\mathbf{x}) = \alpha_F \Delta_f \text{ and } \mathcal{R}_F(\mathbf{x}) = \{r_k(\mathbf{x}) = \gamma_k h_F(\mathbf{x})\}_{k=1}^K.$$

Next, an anisotropic metric along the BRep edge measures whether small gaps exist of the two adjacent faces are kind of "seam edges",

$$d_i(\mathbf{x}) = \|\mathbf{p}_i - \mathbf{x}\|_{A(\mathbf{x})} = \sqrt{(\mathbf{p}_i - \mathbf{x})^\top A(\mathbf{x})(\mathbf{p}_i - \mathbf{x})},$$

$$A(\mathbf{x}) = R(\mathbf{x}) \text{diag}(\sigma_t^{-2}, \sigma_u^{-2}, \sigma_n^{-2}) R(\mathbf{x})^\top.$$

with  $\sigma_t > \sigma_u \geq \sigma_n$  on the edges and  $A(\mathbf{x}) \propto I$  on the BRep faces. Once

#### A.2. SPH kernel with BRep Surface features.

We use the cubic-spline SPH kernel in 3D,

$$W(q, h) = \frac{1}{\pi h^3} \begin{cases} 1 - \frac{3}{2}q^2 + \frac{3}{4}q^3, & 0 \leq q < 1, \\ \frac{1}{4}(2 - q)^3, & 1 \leq q < 2, \\ 0, & q \geq 2, \end{cases}$$

with  $q = d_i(\mathbf{x})/h$ . Non-sharp operational features (filets, chamfers, bosses) receive a curvature-aware smoothing length

$$h(\mathbf{x}) = \eta \min(h_*(\mathbf{x}), R_\kappa(\mathbf{x})),$$

$$R_\kappa(\mathbf{x}) = \frac{1}{\max(|\kappa_1(\mathbf{x})|, |\kappa_2(\mathbf{x})|, \epsilon)}.$$

where  $h_* \in \{h_E, h_F\}$  and  $\kappa_{1,2}$  are the principal curvatures from the B-Rep surface at  $\mathbf{x}$ .

Optionally, a normal-consistency gate improves robustness near joints:

$$g_i(\mathbf{x}) = \exp(-\lambda [1 - |\mathbf{n}(\mathbf{x})^\top \mathbf{n}_i|]),$$

where  $\mathbf{n}_i$  is the scan normal at  $\mathbf{p}_i$  (if available). The anisotropic metric  $A(\mathbf{x})$  and the SPH kernel  $W(q, h)$  jointly ensure coverage of close-proximity vertices and accurate recovery on smooth CAD features without overextending across sharp boundaries, as described in **Eq. 2** of the main matter.

## B. Synthesized Sketch and Realism Analysis

A critical modality in A2Z-10M+ is the *sketch*, which bridges the gap between early-stage conceptual design and precise BRep geometry. Sketches in engineering practice span a wide stylistic spectrum—from quick freehand concept doodles to dimensioned orthographic projections—and a dataset that targets sketch-based CAD retrieval and reconstruction must faithfully cover this range.

### B.1. Generation Pipeline

A2Z adopts a multi-level 3D sketching framework (Sec. 3.3 of the main paper) that builds on recent advances in differentiable stroke rendering. Beginning from the native BRep representation, the pipeline proceeds through four stages:

1. **View capture.** A set of canonical viewpoints is sampled around the CAD model, and perspective/orthographic projections are rendered, preserving the association between projected curves and their parent BRep edges.
2. **Occlusion removal.** Hidden lines are identified via depth-buffer queries and either removed (for engineering-style drawings) or rendered with dashed strokes (for see-through conceptual sketches).
3. **Loop pruning.** Redundant or visually cluttered silhouette loops are simplified while retaining topologically significant contours (e.g., holes, fillets, chamfer boundaries).
4. **Stroke rendering.** The surviving curves are rasterized with stylistic variation (line weight jitter, over-sketching, slight positional noise) to approximate hand-drawn appearance at multiple fidelity levels.

The result is a hierarchy of sketch abstractions—referred to as *Level-2* through *Level-5* in decreasing order of realism—each retaining the semantic edge-type and edge-ID labels inherited from the parent BRep. This label preservation is essential: while pixelated sketches are sufficient for image-based *retrieval*, they cannot drive BRep *reconstruction* without the accompanying semantic edge annotations that specify curve type (line, arc, B-spline), connectivity, and face adjacency.

### B.2. Quantitative Realism Evaluation

To validate sketch realism, we conducted a perceptual evaluation on a held-out set of 5 000 newly generated 2D sketches (half of our evaluation pool, corresponding to ll. 455–464 of the main paper). Each sketch was scored on a 1–10 realism scale by three independent evaluators: two frontier vision-language models (GPT-5, Gemini) acting as automated raters, and a panel of mechanical engineering graduate students serving as human raters.

Table 6 reports the mean and standard deviation per abstraction level.

Several trends emerge. First, *Level-2* sketches, which retain the most geometric detail, receive high scores from automated raters but are judged slightly lower by humans, likely because the density of strokes can appear mechanical rather than hand-drawn. Second, *Level-3* achieves the strongest overall evaluator consensus: it retains enough salient contours for reliable part identification while introducing sufficient stylistic abstraction to appear natural. Third, at the more abstracted *Level-4* and *Level-5*, human raters remain relatively tolerant (scores above 7.6), whereas automated raters penalize the loss of geometric detail more sharply; this gap highlights a known limitation of VLM-based perceptual metrics, which tend to favour high-frequency detail.

These results confirm that the multi-level sketching pipeline produces perceptually convincing outputs across the realism spectrum, supporting both downstream retrieval tasks (where coarse sketches suffice) and reconstruction tasks (where Level-2/3 sketches with full semantic labels are preferred).

## C. Synthesized Scan and Realism Analysis

The second critical modality in A2Z-10M+ is the *synthesized 3D scan*, which simulates the output of real structured-light or laser-based acquisition devices applied to CAD parts. High-fidelity synthetic scans are essential for training and evaluating scan-to-BRep reconstruction methods [5, 7, 10], yet prior datasets typically add only simple Gaussian noise to uniformly sampled point clouds—a poor proxy for the spatially correlated artefacts produced by physical sensors.

### C.1. Artec3D-Calibrated Scan Synthesis

A2Z addresses this gap through a four-stage scan synthesis algorithm (Sec. 3.1 of the main paper) whose parameters are calibrated against ground-truth (GT) scans produced by the Artec3D virtual scanner. Artec3D is a proprietary tool that replicates the optical path, sensor calibration curves, and opto-photonics noise characteristics of its commercial structured-light hardware; scans produced by this tool are therefore treated as *real ground truth* for evaluation purposes. The four stages introduce progressively more realistic artefacts:

1. **Step I (Uniform sampling & base noise).** Points are sampled uniformly on the BRep surface and perturbed with isotropic Gaussian noise calibrated to the Artec3D sensor’s baseline repeatability specification.
2. **Step II (View-dependent density).** A virtual sensor viewpoint is introduced, and point density is modulated by the local surface-to-sensor angle, replicating the reduced sampling density observed on oblique



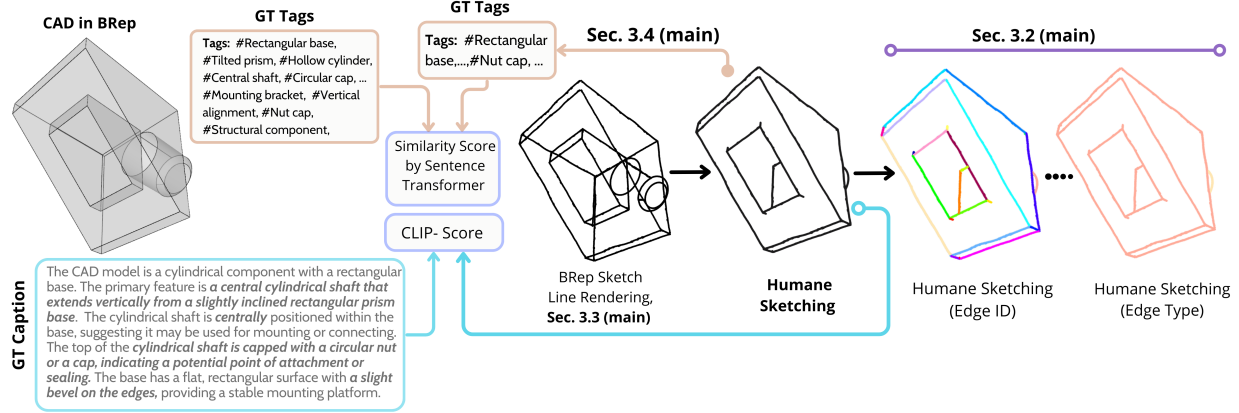


Figure 1. **Realism of Sketch.**: We compute a CLIP cosine-similarity score of **30.58** between generated captions and the corresponding 2D sketches, confirming that textual descriptions are well-grounded in visual content. Additionally, a tag-level similarity score of **64.7%** (CAD-derived tags vs. sketch-derived tags) demonstrates that the hierarchical tag vocabulary remains consistent across the BRep and sketch modalities.

Evaluator	Modality	Level-2	Level-3	Level-4	Level-5
GPT-5	sketch	$7.65 \pm 0.34$	$7.20 \pm 0.48$	$5.53 \pm 0.38$	$5.05 \pm 0.39$
Gemini	sketch	<b><math>9.01 \pm 0.28</math></b>	$7.33 \pm 0.52$	$6.06 \pm 0.58$	$3.92 \pm 0.47$
Human	sketch	$7.20 \pm 0.32$	<b><math>7.86 \pm 0.38</math></b>	$7.85 \pm 0.44$	$7.67 \pm 0.71$

Table 1. Realism scores (mean  $\pm$  std, 1–10 scale) for *multi-level 2D sketches* evaluated by automated VLM raters and human engineering students. Higher is better. Level-3 achieves the strongest evaluator consensus, balancing abstraction with recognizability.

and self-occluded regions.

- Step III (Structured artefacts).** Spatially correlated noise patterns—edge ringing, specular dropout, and inter-reflection ghosts—are layered according to material BRDF models and local curvature, emulating systematic scanner errors absent from i.i.d. noise models.
- Step IV (Full sensor simulation).** The complete Artec3D opto-photonic pipeline is applied, including lens distortion, structured-light fringe decoding residuals, and outlier generation at depth discontinuities, yielding scans that closely match real GT.

## C.2. Stage-Wise Realism Evaluation

To quantify the realism gain at each stage, we synthesized scans from CC3D CAD models whose Artec3D GT scans are available. We evaluated each stage using the same three-evaluator protocol described in Sec. B.2, *i.e.*, automated VLM raters and human experts. The Table 7 reports the results employing the visual comparison tool shown in Fig. 3.

The results reveal a clear monotonic improvement in perceived realism across stages. Steps I and II, which model only point distribution and view-dependent density, receive modest scores (6.3–7.9), indicating that

density variation alone is insufficient to fool either human or automated raters. The introduction of structured artefacts in Step III yields a noticeable jump (roughly +1 point across all evaluators), confirming that spatially correlated sensor errors are a primary perceptual cue for scan authenticity. Finally, Step IV achieves near-ceiling scores— $9.83 \pm 0.22$  from human raters—demonstrating that the full opto-photonic simulation produces scans virtually indistinguishable from real Artec3D acquisitions.

The extremely low variance at Step IV for human raters ( $\sigma = 0.22$ ) is particularly noteworthy: it indicates strong inter-annotator agreement that the synthesized scans are realistic, providing a high-confidence quality floor for downstream learning tasks such as boundary detection, surface segmentation, and BRep fitting.

## D. Prompt for CAD Description and Tags

We have structured our prompt template for a mixture of VLMs – Qwen3-30B [11] and InternVL3 [1] (see Section 3.4 of the main matter)—and defined a consistent textual description and tag set. In the prompt template, the following segments are present:

Evaluator	Modality	Step-I	Step-II	Step-III	Step-IV
GPT-5	scan	$6.51 \pm 0.84$	$6.54 \pm 1.01$	$7.58 \pm 0.58$	<b><math>9.41 \pm 0.60</math></b>
Gemini	scan	$7.73 \pm 1.48$	$7.91 \pm 0.92$	$8.54 \pm 0.78$	<b><math>9.28 \pm 0.81</math></b>
Human	scan	$6.32 \pm 0.29$	$6.58 \pm 0.80$	$7.20 \pm 0.51$	<b><math>9.83 \pm 0.22</math></b>

Table 2. Realism scores (mean  $\pm$  std, 1–10 scale) for *synthesized scans* at each pipeline stage, evaluated against Artec3D ground-truth scans of CC3D models. Refer to Figure. 5 for visuals. **Step IV** achieves near-perfect human consensus.

#### Overall LLM’s Role

“You are a senior mechanical design engineer with deep expertise in component identification and sourcing. You are shown 12 images of a single CAD model from multiple views, arranged in a grid. Your task is to analyse the geometry and produce a technical description and classification tags suitable for a professional engineering database”.

After defining the role, we fix the rules for the description in in-context prompt learning in three different parts: (1) Description, (2) Filtering and Constraints, and (3) Desired Output Structure.

### D.1. Description

Write a single, concise paragraph (maximum 150 words) that follows these guidelines:

1. **Primary identification.** Use specific, standard mechanical terminology only when the geometry makes the class obvious; if the type is unclear, describe the geometry directly and do *not* speculate about its industry role.
2. **Justification and geometry.** Justify your identification by describing the key geometric features that support it. If you cannot confidently name the part, instead of describing its dominant geometric forms (cylinders, prisms, ...) and their spatial relationships using CAD terms (concentric, axial, radial, ...).
3. **Feature details.** Mention mounting features, holes, cut-outs, bosses, slots, or surface transitions (fillets, chamfers, ...) that appear critical to the part’s function.
4. **Synthesis of views.** Ensure that the description is a cohesive synthesis of all views, explicitly noting overall symmetries and important spatial relationships.
5. **Potential application:** Briefly state a plausible industrial application or use case, and the likely engineering function of the component based on its geometry.

### D.2. Applying Filtering and Constraints

#### Filtering and Constraints

We have implemented some restrictions and ethical filtering constraints for Qwen3-30B [11] to generate the final response. For instance,

1. Do **not** use canned template phrases such as “This appears to be a ... shaft coupling” or “The component is a ... cylinder” unless the object is genuinely that specific type. Be original in your wording.
2. The final tag list must contain 20 items in a flat structure. Do not include headings like “Category/Class:” in the output.
3. Avoid generic, overused terms (e.g., *flange*, *bracket*, *housing*) unless the geometry explicitly and unambiguously matches those features, and you are more than 90% confident.

### D.3. Output Format

The response must follow this exact structure:

#### Output Structure

```
Description: {Your paragraph here}
Tags:
- Tag 1
- Tag 2
- ...
- Tag 10
```

## E. Hierarchical Tags for CAD Retrieval

Geometric feature recognition [4, 9] on BRep, and then streamlining those features to replace traditional part modeling techniques [3], can transform complex CAD retrieval for reuse purposes in design. Like many other important downstream tasks of BRep learning the retrieval and identification of CAD models by their geometric features [6] automate manufacturing steps by saving time and reducing human error. For this purpose, a well-structured database of part components is necessary for ingestion. Such a database serves as a go-to reference that links the Computer-Aided Design (CAD)

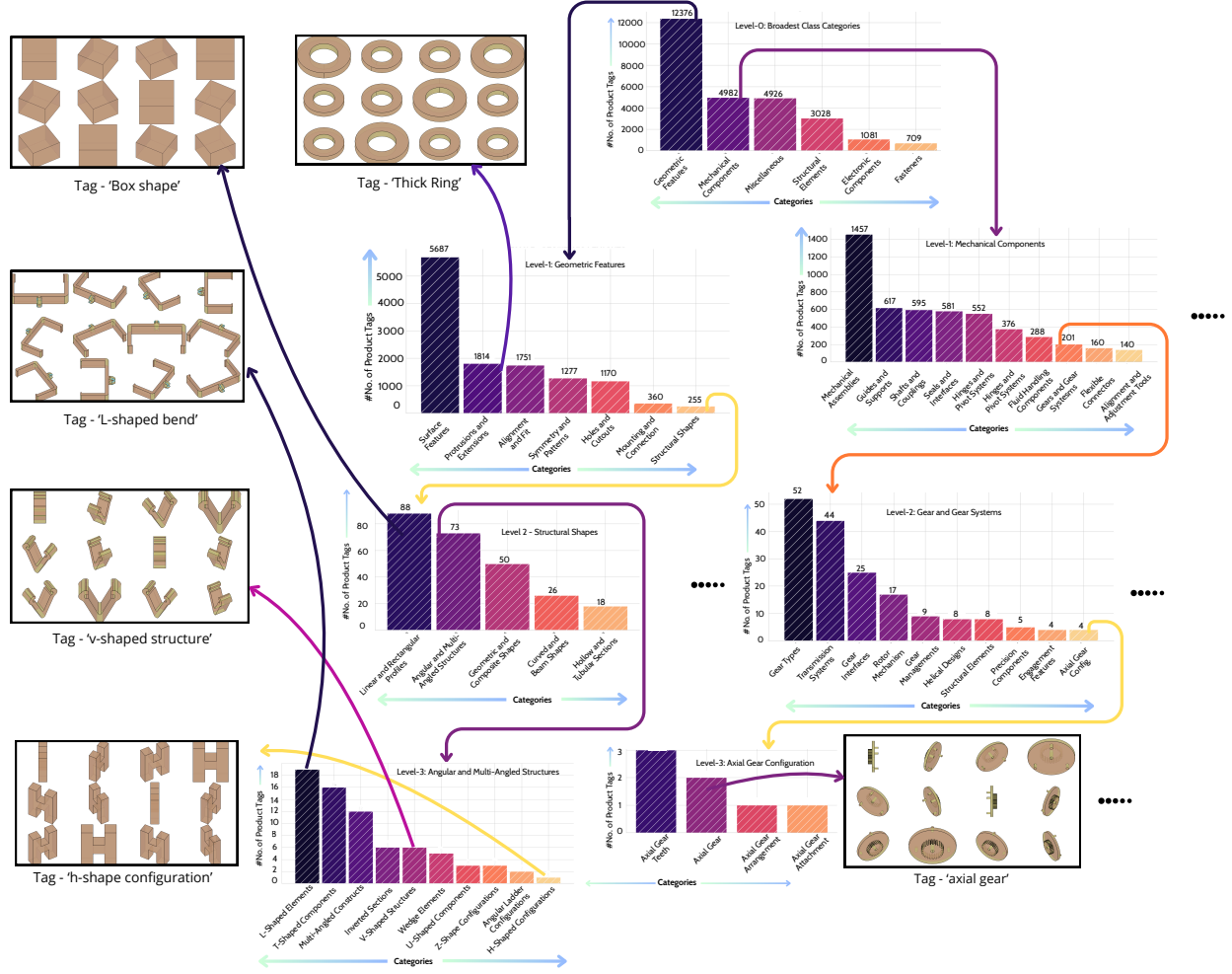


Figure 2. Hierarchical categorization of 1 million CAD models for retrieval and causal mining. The two examples illustrate how – (*on the right*) one can fetch some CAD models when asked to find – ‘A **mechanical component** associated to **Gear type transmission** systems supporting **Axial** rotation’.

process to the Computer-Aided Manufacturing (CAM) workflow. Our methodology for this hierarchical mining and retrieval of CAD based on their captions and unique tags is described below:

### E.1. Semantic Hierarchy over 1M A2Z Models

We begin with a corpus of roughly one million CAD models from the ABC dataset, each annotated with a visual caption and up to ten free-form tags, as mentioned in the main matter. An initial test run is first made on the  $N = 27,102$  distinct tags that appear throughout the corpus. We construct a semantic hierarchy that simultaneously reflects human design intuition and the empirical distribution of the data. We iteratively use Qwen3–30B [11] as an LLM to serve as a semantic classifier, running over  $\lfloor \frac{N}{100} \rfloor$  unique tags in a chunk and then clustering them consistently across other chunks. Therefore, we have created semantic clustering with a maximum depth of 5 in distinct phases.

**Phase-1: high-level assignment.** Every tag  $t$  is first routed into one of  $K$  coarse parent categories  $\{c_1, \dots, c_K\}$  (e.g., *Geometric Features*, *Mechanical Components*, *Structural Elements*, *Electronic Components*, *Fasteners*, plus an *Unclassified* bucket). The model is prompted with the tag in isolation and returns a single category label, yielding a mapping

$$\phi_1 : t \mapsto c_k.$$

This step ensures that all later reasoning is constrained within a semantically coherent region of the space.

**Phase-2: recursive refinement.** Within each coarse category, we recursively subdivide tag sets into more specific groups. Given a set  $S \subset \mathcal{T}$  of tags that share the same parent, the LLM is asked to propose between five and ten meaningful sub-groups that partition  $S$  (e.g., under *Geometric Features*, we obtain groups such as *box-like shapes*, *L-bends*, *V-shaped structures*, and *axial gears*, as shown in Fig. 2). The process is repeated for

Table 3. Distribution of the 27,102 unique tags over top-level categories and their Percentage (%) of CAD models.

Category	Tag Count	Percentage
Geometric Features	12,230	45.13%
Unclassified	6,068	22.39%
Mechanical Components	4,601	16.98%
Structural Elements	2,724	10.05%
Electronic Components	801	2.96%
Fasteners	678	2.50%

each child group until either (i) the group contains fewer than ten tags or (ii) a maximum depth is reached. The result is an irregular, data-driven tree that is deep where the vocabulary is rich and shallow where it is sparse.

**Phase-3: attaching model identifiers.** After the structure of the tree is fixed, we traverse it and for each leaf node  $\nu$  we attach the set of CAD models

$$\mathcal{M}_\nu = \{m \in \mathcal{M} \mid \text{tag}(m) \ni t \text{ and } t \in \nu\}$$

that carry any of the tags stored at  $\nu$ . As a consequence, every model is reachable from multiple semantic leaves, coupling fine-grained vocabulary (e.g., “*axial gear with keyway*”) with the raw geometry of the associated CAD.

**Phase 4: statistical reporting.** Finally, the fully assembled hierarchy is summarized to support downstream analysis and dataset curation. Table 3 reports the observed distribution of tags over the root categories, revealing, for instance, that nearly half of all tags are geometric in nature, while mechanical and structural terms jointly account for roughly one quarter of the vocabulary. These statistics are later used to rebalance training sets and design retrieval strategies that avoid overfitting to overrepresented classes.

## E.2. Heterogeneous Graph Representation

The hierarchical taxonomy is converted into a heterogeneous graph that unifies tags, categories, captions, and CAD models. Let  $\mathcal{M}$  denote the set of model nodes,  $\mathcal{T}$  the set of tag nodes, and  $\mathcal{C}$  the set of category nodes (internal and root). The vertex set is

$$V = \mathcal{M} \cup \mathcal{T} \cup \mathcal{C},$$

and the edge set  $E$  contains four main relation types:

1. **Taxonomic edges** ( $c_{\text{parent}}, c_{\text{child}}$ ) encode the “is-a” relation of the hierarchy tree.
2. **Tag–category edges** ( $t, c$ ) for every tag  $t$  assigned to a leaf category  $c$ .
3. **Model–tag edges** ( $m, t$ ) whenever the model  $m$  carries the tag  $t$ .
4. **Co-occurrence edges** ( $t_i, t_j$ ) with weight  $w_{ij}$  proportional to the point-wise mutual information or normalized co-frequency of  $t_i$  and  $t_j$  across models.

In the future, we aim to equip each node with learned feature vectors, *i.e.*, text embeddings generated using their captions passed through a foundational model like GPT[8]. The resulting structure is a richly typed knowledge graph that captures both symbolic semantics (“V-shaped structural bracket”) and descriptions.

## E.3. Graph Mining and Retrieval Mechanism

Given this representation, both discovery and retrieval reduce to structured navigation over the graph. **Mining semantic and geometric patterns.** We apply community detection and frequent subgraph mining to the tag-tag and model-tag subgraphs to uncover recurrent design motifs. For example, high-weight cliques connecting *box shape*, *rectangular pocket*, and *through hole* correspond to machining templates frequently seen in prismatic parts, while another community might link *axial gear*, *shaft*, and *keyway*. At higher levels, modularity-based clustering over the category graph yields the thematic groupings visualized in the first image, where related subtrees (e.g., gear systems, bent profiles, and multi-angled structures) are summarized as coherent “families” of shapes. These mined structures directly inform both dataset balancing and the design of tag suggestions for new models.

## F. GPT-5, Gemini, and Human Evaluation.

In continuation of *Section 5* of the main matter, the Fig. 3 describes how a multi-stage workflow was designed to measure the quality of **Face-Type**, **Face-ID**, **Boundary-Type**, and **Boundary-ID** annotations on both 3D scans and their corresponding B-Rep visualizations. As illustrated in the evaluation diagram, the randomly selected 10K samples are first divided into multiple “Eval Batches”, each containing 10 CAD models. For every batch, the 3D scans were color-mapped using the BRep entity feature (see the color map in the teaser figure of the main matter). Thereafter, a corresponding BRep model is also visualized using OpenCASCADE, where the geometric entities are color-mapped to serve as ground truth. Automatic quick snapshots are captured when the CAD models and 3D Scans are collectively rotated to other views. These stacked images are passed to Gemini, GPT-5 to collect the annotation quality scores. The human evaluators also submit their scores in parallel. These scores are aggregated into structured matrices—one for each type of BRep entity labels on the right side of Fig. 3. To avoid redundancy, the other similar matrices for sketches and boundaries are not illustrated. However, we have conducted a similar evaluation for all the boundary, corner, and face-related BRep entities.

## F.1. Chunk-wise Evaluation of Model

We have also unrolled the chunk-wise (0–100) recall and precision measurements for the boundary and junction



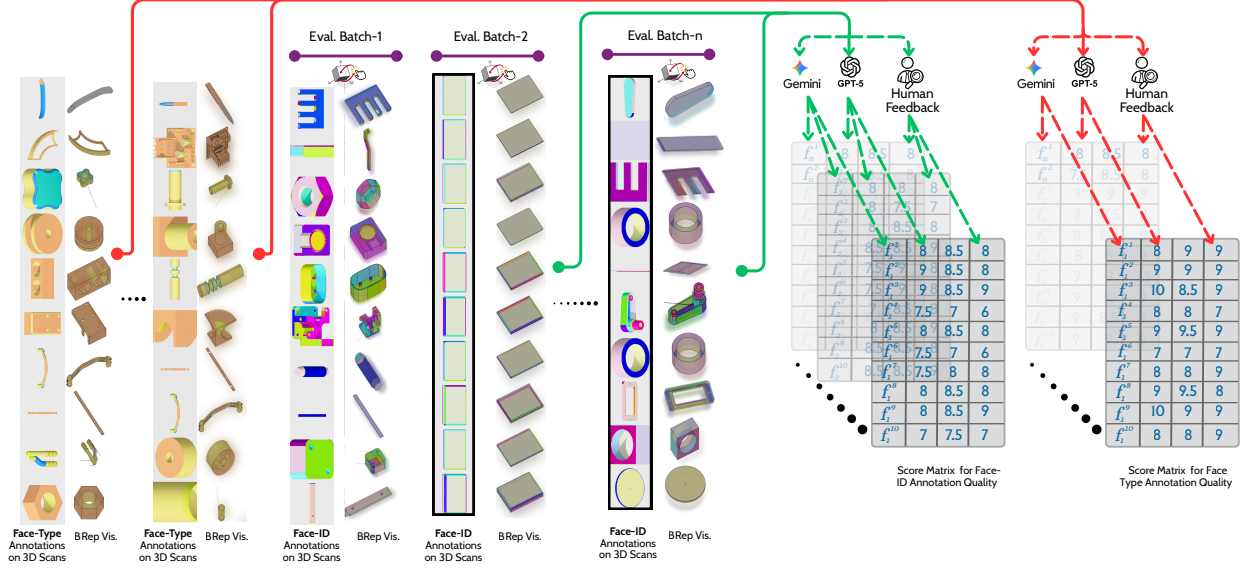


Figure 3. This figure visually describes how evaluation and scoring mechanisms from GPT-5, Gemini, and the Human feedback-based evaluation process for different types of annotations on 10K randomly selected samples.

tasks in Tab. 4. The green colored region, *i.e.*, chunks 70–99, contains 3D scans of a very high-resolution (16x upsampled version of ABC meshes). Therefore, the recall-precision for those chunks is consistently strong. These unrolled results complement the average metrics reported in the main paper. The Fig. 4 illustrated very high-quality boundary and corner detection on a randomly downloaded 3D scan file from the online GrabCAD (<https://grabcad.com/library>) library

## G. More Examples of Caption and Tags

Figs. 6 to 8 illustrates many examples of CAD captions and tags fetched randomly from our **A2Z**. The example captions and tags reveal a *geometry-first* reasoning style: each description starts by carefully characterizing observable shapes (cylinders, plates, frames, blades) and spatial relations (radial, concentric, open-frame), and only then moves to *functional hypotheses* such as cooling, support, or grip. Across both mechanical and decorative objects, the text consistently integrates *local features* (holes, flanges, chamfers, cut-outs) with *global structure* (symmetry, layering, pose) to justify why a part might behave as a fan housing, a probe, a bracket-like stand, or a stylized human figure. The sets of tags act as compact *multi-axis descriptors*, jointly encoding category (e.g., component vs. figurine), function (cooling, fastening, display), dominant geometry (cylindrical body, radial blades, open frame), and likely application domain (industrial machinery, ventilation, artistic display). Together, these captions and tags show how language can remain tightly grounded in CAD evidence while still offering informative, searchable abstractions

that bridge *shape, function, and use context*.

## H. More Examples of Electronic Enclosures

The Fig. 9 and Fig. 10 demonstrate samples that were randomly drawn from our newly added CAD models in **A2Z**, consisting of charging ports, usb-a/b/c type ports, and other socket-type enclosures. The present CAD market for electronic equipment is larger than that for mechanical engineering. The newly added CAD models and results from our foundational model of boundary and junction detection will be highly effective in standardizing the BRep entities of such complex CAD models. Unlike the charging port design data, where the complexity of CAD models is greater in terms of the topology of the folding faces/boundaries, the Fig. 11 illustrates a different challenge regarding the geometric continuity of the BRep faces/boundaries around the corner region. This type of data annotation and our foundational model for parsing the BRep entities are close to industrial settings for accurate CAD reconstruction.

## I. Examples of Annotations and Predictions

In Fig. 12 we add more examples of our rich multi-modal annotations on 3D scans and sketches. The color coding is shown in Fig. 3 of the main matter.

## References

- [1] Zhe Chen, Weiyun Wang, Hao Tian, Shenglong Ye, Zhangwei Gao, Erfei Cui, Wenwen Tong, Kongzhi Hu, Jiapeng Luo, Zheng Ma, et al. How far are we to gpt-4v? closing the gap to commercial multimodal models with open-source suites. *Science China Information Sciences*, 67(12):220101, 2024. 3



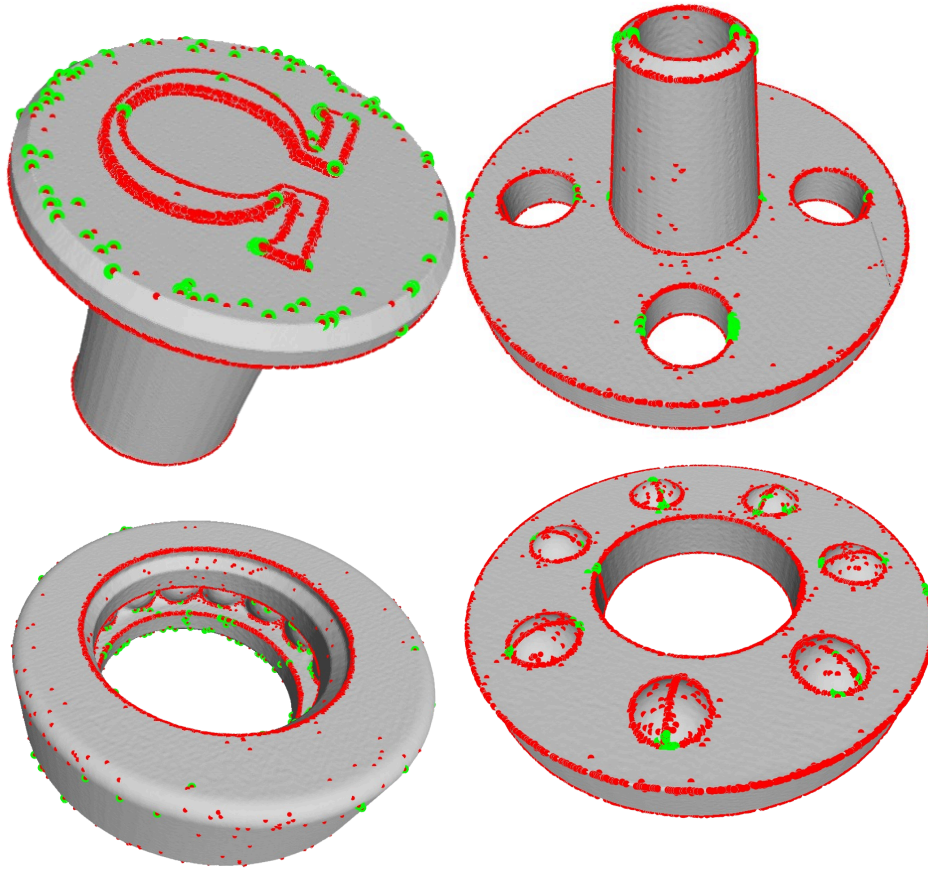


Figure 4. Boundary and Corner detection on randomly selected real-world 3D scan data from the Grab-CAD repository.

- [2] Kseniya Cherenkova, Djamila Aouada, and Gleb Gusev. Pvdconv: Point-voxel deconvolution for autoencoding cad construction in 3d. pages 2741–2745, 2020. [9](#)
- [3] John G Cherng, Xin-Yu Shao, Yubao Chen, and Peter R Sferro. Feature-based part modeling and process planning for rapid response manufacturing. *Computers & industrial engineering*, 34(2):515–530, 1998. [4](#)
- [4] Yongkang Dai, Xiaoshui Huang, Yunpeng Bai, Hao Guo, Hongping Gan, Ling Yang, and Yilei Shi. Brepformer: Transformer-based b-rep geometric feature recognition. In *Proceedings of the 2025 International Conference on Multimedia Retrieval*, page 155–163, New York, NY, USA, 2025. Association for Computing Machinery. [4](#)
- [5] Haoxiang Guo, Shilin Liu, Hao Pan, Yang Liu, Xin Tong, and Baining Guo. Complexgen: Cad reconstruction by b-rep chain complex generation. *ACM Transactions on Graphics (TOG)*, 2022. [2](#)
- [6] Jinwon Lee, Changmo Yeo, Sang-Uk Cheon, Jun Hwan Park, and Duhwan Mun. Brep gat: Graph neural network to segment machining feature faces in a b-rep model. *Journal of Computational Design and Engineering*, 10(6):2384–2400, 2023. [4](#)
- [7] Yujia Liu, Anton Obukhov, Jan Dirk Wegner, and Konrad Schindler. Point2cad: Reverse engineering cad models from 3d point clouds. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 3763–3772, 2024. [2](#)
- [8] Josh OpenAI, Achiam, Steven Adler, Sandhini Agarwal, Lama Ahmad, Ilge Akkaya, Florencia Leoni Aleman, Diogo Almeida, Janko Altschmidt, Sam Altman, Shyamal Anadkat, et al. Gpt-4 technical report. *arXiv preprint arXiv:2303.08774*, 2023. [6](#)
- [9] Bonsa Regassa Hunde and Abraham Debebe Woldeyohannes. Future prospects of computer-aided design (cad) – a review from the perspective of artificial intelligence (ai), extended reality, and 3d printing. *Results in Engineering*, 14:100478, 2022. [4](#)
- [10] Cheng Wang, Xinzhu Ma, Bin Wang, Shixiang Tang, Yuan Meng, and Ping Jiang. Point2Primitive: CAD Reconstruction from Point Cloud by Direct Primitive Prediction. *arXiv e-prints*, art. arXiv:2505.02043, 2025. [2](#)
- [11] An Yang, Anfeng Li, Baosong Yang, Beichen Zhang, Binyuan Hui, Bo Zheng, Bowen Yu, Chang Gao, Chengen Huang, Chenxu Lv, et al. Qwen3 technical report. *arXiv preprint arXiv:2505.09388*, 2025. [3](#), [4](#), [5](#)



Figure 5. Visual comparison between CC3D [2] CAD models, its real scan pairs, and **synthesized scans from CAD (ours)**

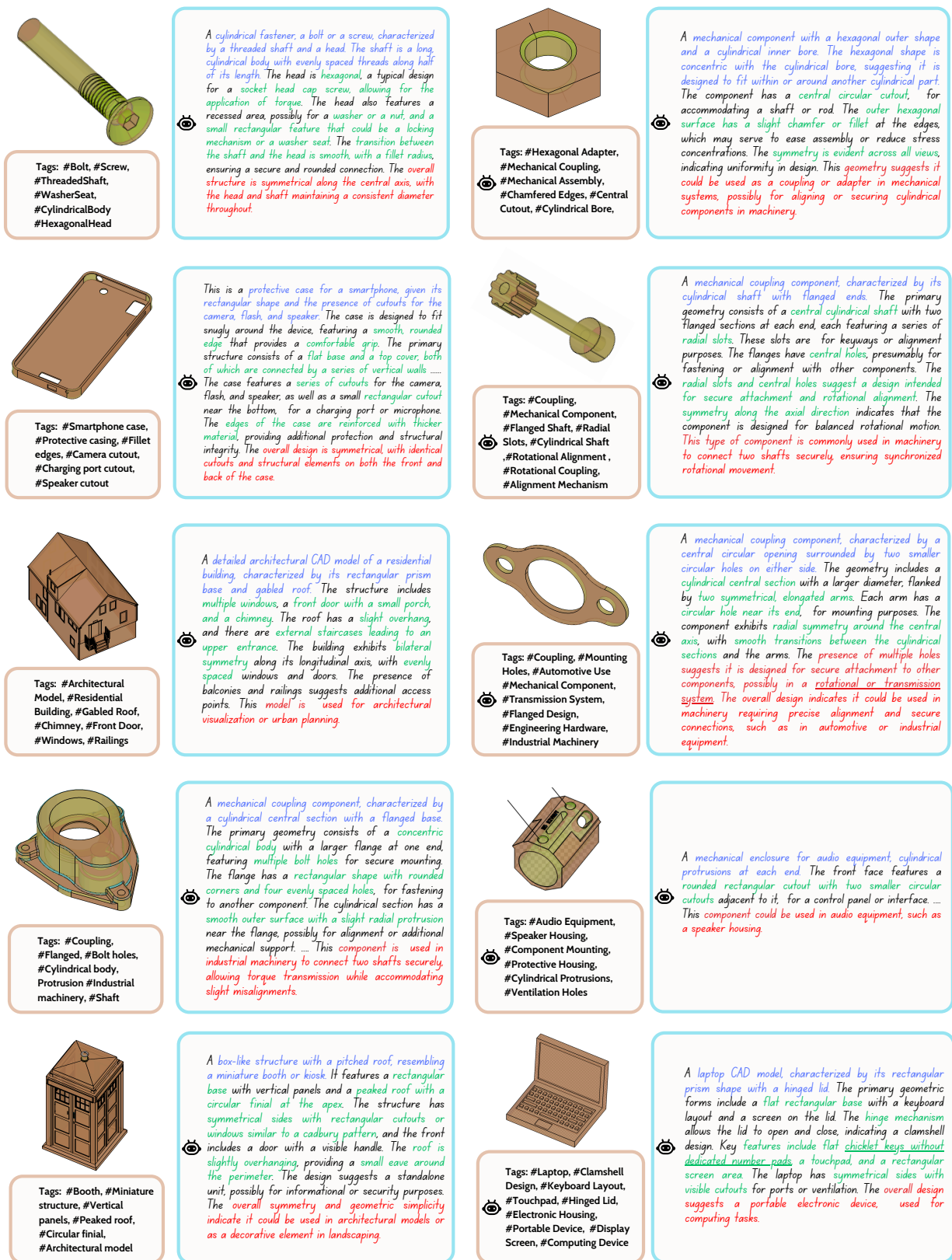
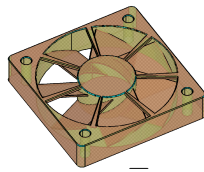


Figure 6. Visual elaboration of the textual captions on more complex CAD models.



Figure 7. Visual elaboration of the textual captions on more complex CAD models





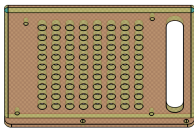
Tags: #Cooling system, #Radial blades, #Axial flow, #Radial fan, #Blade assembly, #Ventilation component, #Airflow

A mechanical component with a central hub and multiple radial blades extending outward, enclosed within a square frame. The central hub is circular and concentric with the radial blades, which are evenly spaced and symmetrically arranged. The square frame has four mounting holes at each corner, suggesting it is designed for secure attachment to another surface. The component exhibits a combination of cylindrical and prismatic forms, with the blades creating a fan-like structure. The overall symmetry and the presence of mounting holes indicate it is used for ventilation or cooling purposes, as a fan housing or blower assembly.



Tags: #Trophy, #Award Cup, #Grip Feature, #Ceremonial Use, #Decorative Item, #Competition Award, #Angular Handles, #Flared Top, #Cylindrical Body, #Industrial Design

A trophy or an award cup, characterized by its symmetrical, three-dimensional geometry. It consists of a cylindrical body with a flared top, supported by a rectangular base. The cylindrical body has two handles extending radially from the upper section, which are angular and slightly curved. The base is square with a central circular platform that supports the cylindrical body. The spatial relationship between the components is concentric, with the cylindrical body centered on the base. The handles provide a grip feature, while the cylindrical body and base are connected seamlessly, suggesting a monolithic structure. This design is typical for decorative or ceremonial objects, used in events or competitions for recognition purposes.



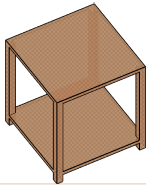
Tags: #Heat sink, #Metal plate, #Ventilation, #Circular holes, #Grid pattern, #Uniform distribution, #Plate

A rectangular plate with a series of circular holes arranged in a 4x8 grid pattern. The plate features a rectangular cutout on one side for mounting or alignment purposes. The holes are evenly spaced and cover most of the plate's surface, which resembles an optical breadboard. The edges of the plate are smooth, with no visible chamfers or fillets, indicating a straightforward manufacturing process. The symmetry of the hole pattern and the rectangular cutout suggests that this part could be used in applications requiring uniform distribution of forces or airflow. The overall geometry and features imply that it might serve as a mounting plate or a heat sink in industrial machinery.



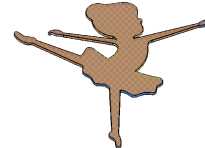
Tags: #Floral Motif, #Display Item, #Decorative Box, #Leaf Pattern, #Ornamental Artwork, #Wood-Like Material, #Handcrafted Appearance, #Artistic Embellishment

A rectangular box with intricate surface engravings. It has a symmetrical design with a central ornamental pattern featuring floral and leaf motifs. The box has a flat top surface with detailed embossing, and the sides are slightly raised, forming a shallow rectangular prism. The corners are rounded, and there are no visible mounting features, holes, or cutouts. The overall symmetry and decorative nature suggest it could be a decorative box or container, used for aesthetic purposes or as a small storage unit.



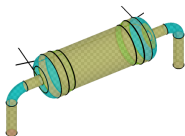
Tags: #Shelving Unit, #Structural Support, #Support Frame, #Open Frame, #Display Stand, #Industrial Furniture

A two-tiered open-frame structure with a rectangular base and an elevated rectangular platform. The primary geometric forms include rectangular prisms, with the upper and lower tiers aligned axially. The design features clean, sharp edges with no visible fillets or chamfers. The structure lacks any mounting holes or cutouts, suggesting it may serve as a standalone unit or a part of a larger assembly. The symmetrical arrangement and open design indicate potential use as a shelving unit or a support frame. The simplicity and open nature suggest it could be used in storage solutions or as a display stand in industrial or office settings.



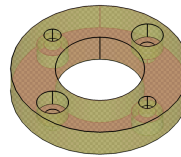
Tags: #Human Figure, #Stylized Representation, #Ballet Pose, #Human Movement, #Figure Sculpture, #Artistic Model, #Dance Study, #Pose Variation

A stylized representation of a human figure in various ballet poses, rather than a mechanical component. The geometry consists of a central body with limbs extended in different directions, suggesting movement and flexibility. The figure has a rounded head, a torso with a flared skirt-like structure, and elongated limbs. The limbs are positioned in various ballet stances, indicating dynamic poses. There are no visible mounting features, holes, or cutouts, and the surface transitions are smooth with no distinct fillets or chamfers. The overall symmetry is bilateral, with each pose reflecting a mirrored counterpart. While not a mechanical part, this figure could be used in artistic or educational applications, as a model for studying human movement or dance.



Tags: #90-Degree Pipe Bends, #Piping, #Pipe Fitting, #Flanged Adapter, #Cylindrical Component, #Gas Transfer System

A cylindrical flow-through canister with pipe connections. It features a central axis-symmetric cylinder with stepped collars/weld bands near each end, each end pipe with a smooth 90-degree elbow bend. The body is adorned with several bands, for reinforcement or attachment purposes. The component exhibits axial symmetry, with the bands evenly spaced around the circumference. The flanges at the ends suggest it is designed for secure mounting or connection to other components. The presence of these flanges and the overall cylindrical shape indicates that it could be used in fluid or gas transfer systems, as a coupling or adapter, and with constant-ID tubing for uninterrupted flow.



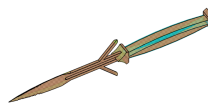
Tags: #Flanged Coupling, #Spacer, #Mounting Holes, #Industrial Machinery, #Mechanical Assembly, #Central Bore, #Fastening Feature, #Radial Holes, #Shaft Alignment

A circular mechanical component with a central hole and multiple peripheral features. It is primarily cylindrical with a concentric arrangement of holes around its perimeter. The component has a central circular cutout, surrounded by a thicker ring. There are eight evenly spaced holes around the outer ring, each with a smaller concentric hole, for mounting or fastening purposes. The inner and outer edges of the component have slight chamfers, and the overall design suggests symmetry and balance. This geometry is typical of a flanged coupling or a spacer used in mechanical assemblies to connect shafts or provide spacing between components. The mounting holes indicate it is designed to be secured in place, in machinery or industrial equipment.



Tags: #Probe, #Stylus, #3D Scanning, #Conical Tip, #Measurement Tool, #Contact Measurement, #Precision Component

A cylindrical mechanical component with a tapered tip, a type of probe or stylus. The primary geometry consists of a long, straight cylindrical body with a smaller diameter section near the top, transitioning into a conical tip. The cylindrical body has a uniform diameter, while the conical tip is sharply pointed, suggesting precision in its design. The symmetry is axial, with the conical tip aligned along the central axis of the cylinder. This geometry suggests that the component could be used in applications requiring precise contact or measurement, such as in sensors, probes, or styluses for 3D scanning.



Tags: #Spear, #Javelin, #Pointed tip, #Grip section, #Aerodynamic shape, #Parallel prongs, #Throwing spear, #Combat tool, #Crossguard, #Pointed projectile

A spear or javelin, characterized by a long, slender shaft with a pointed tip. The shaft is cylindrical with a consistent diameter, transitioning into a tapered point at one end. The other end features a crossguard, which is perpendicular to the shaft, providing balance and protection. The crossguard is composed of two parallel prongs, symmetrically aligned. The spear also includes a grip section, slightly thicker than the main shaft, designed for secure handling. The overall design suggests a focus on aerodynamic efficiency and stability during flight, with the crossguard enhancing control and safety.

Figure 8. Visual elaboration of the textual captions on more complex CAD models



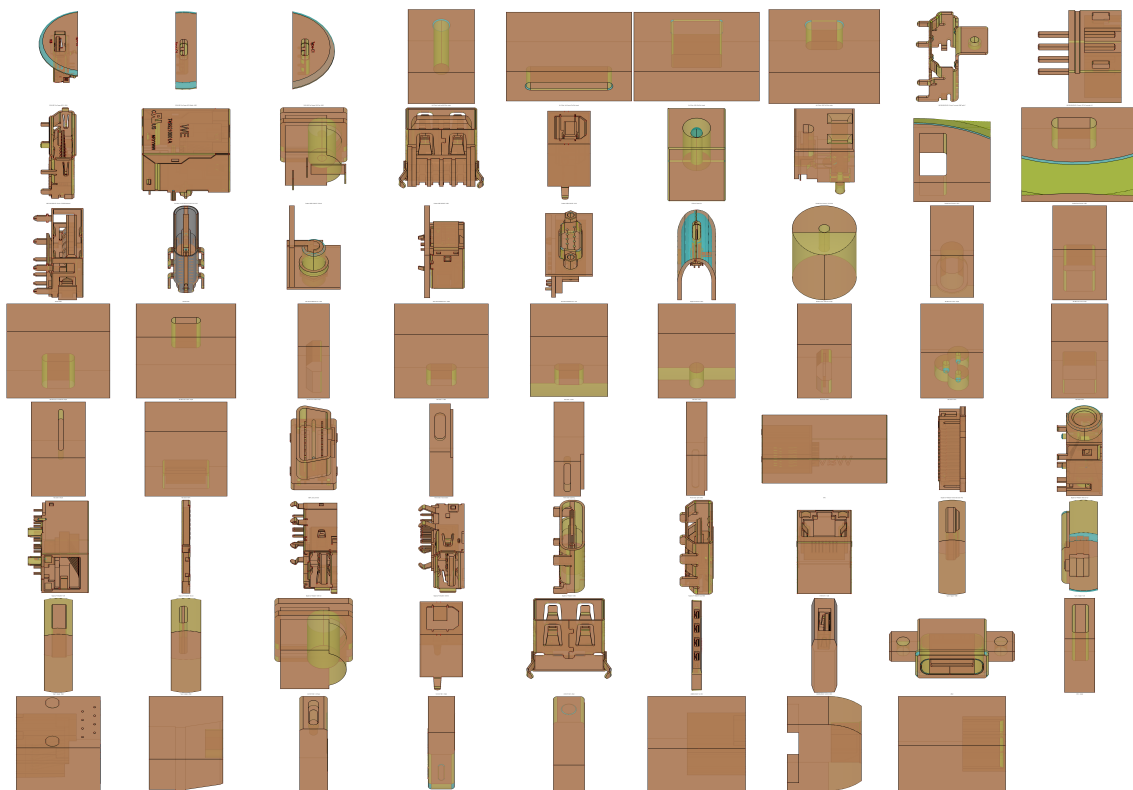


Figure 9. Visuals of some randomly selected CAD models of charging ports under **A2Z**. Color codes define face-Types.

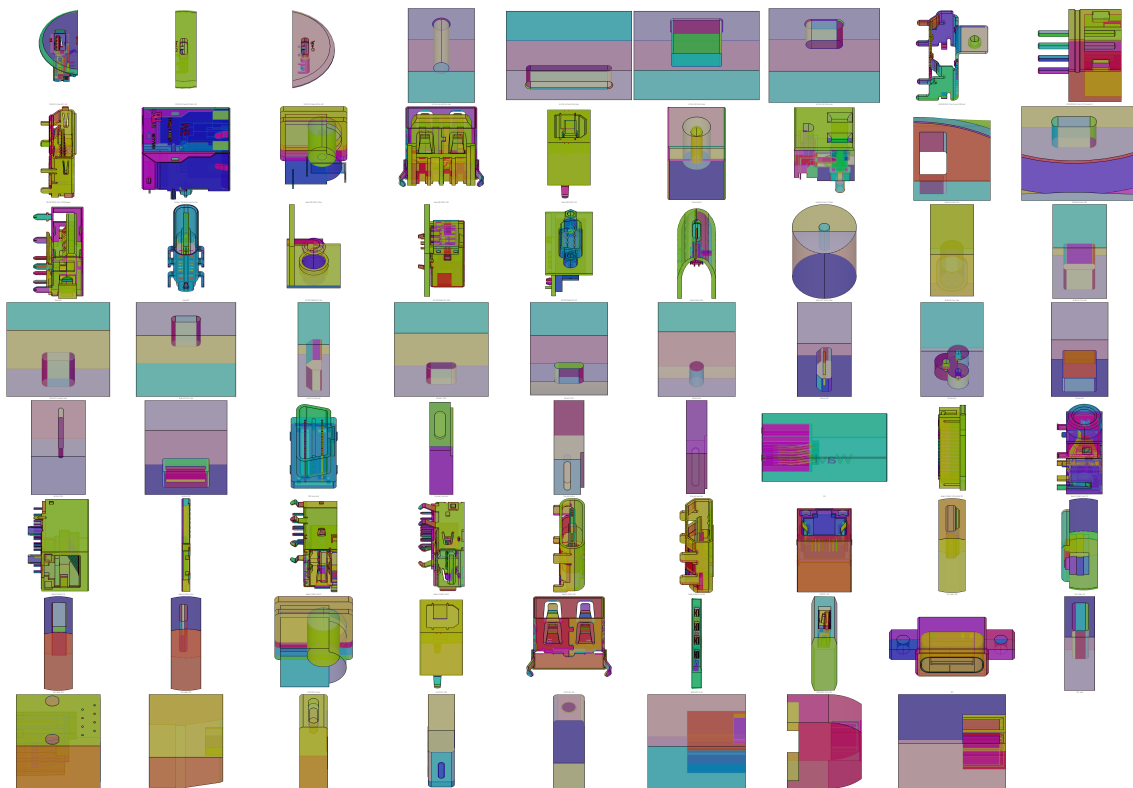


Figure 10. Visuals of same CAD models of charging ports under **A2Z** where color codes define face-IDs.

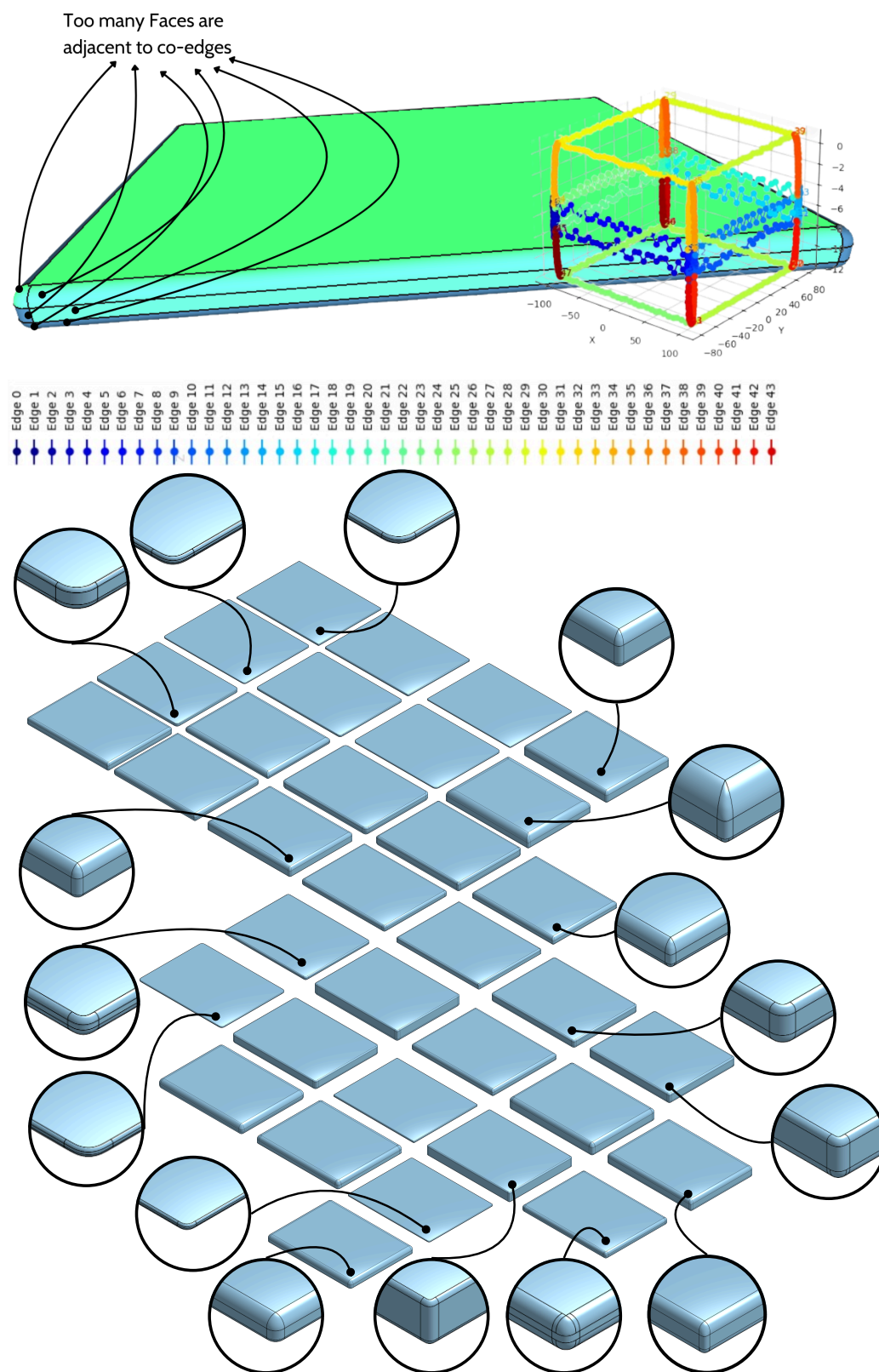


Figure 11. Visuals of some randomly selected samples of CAD models for Tablet type enclosures (from Approx. 20K Electronic enclosure models) with varying corner geometry and topology (*at bottom*). A closer look at the BRep faces, edges, and vertices near the corner parts (*on top*) shows how many faces are adjacent to each edge.

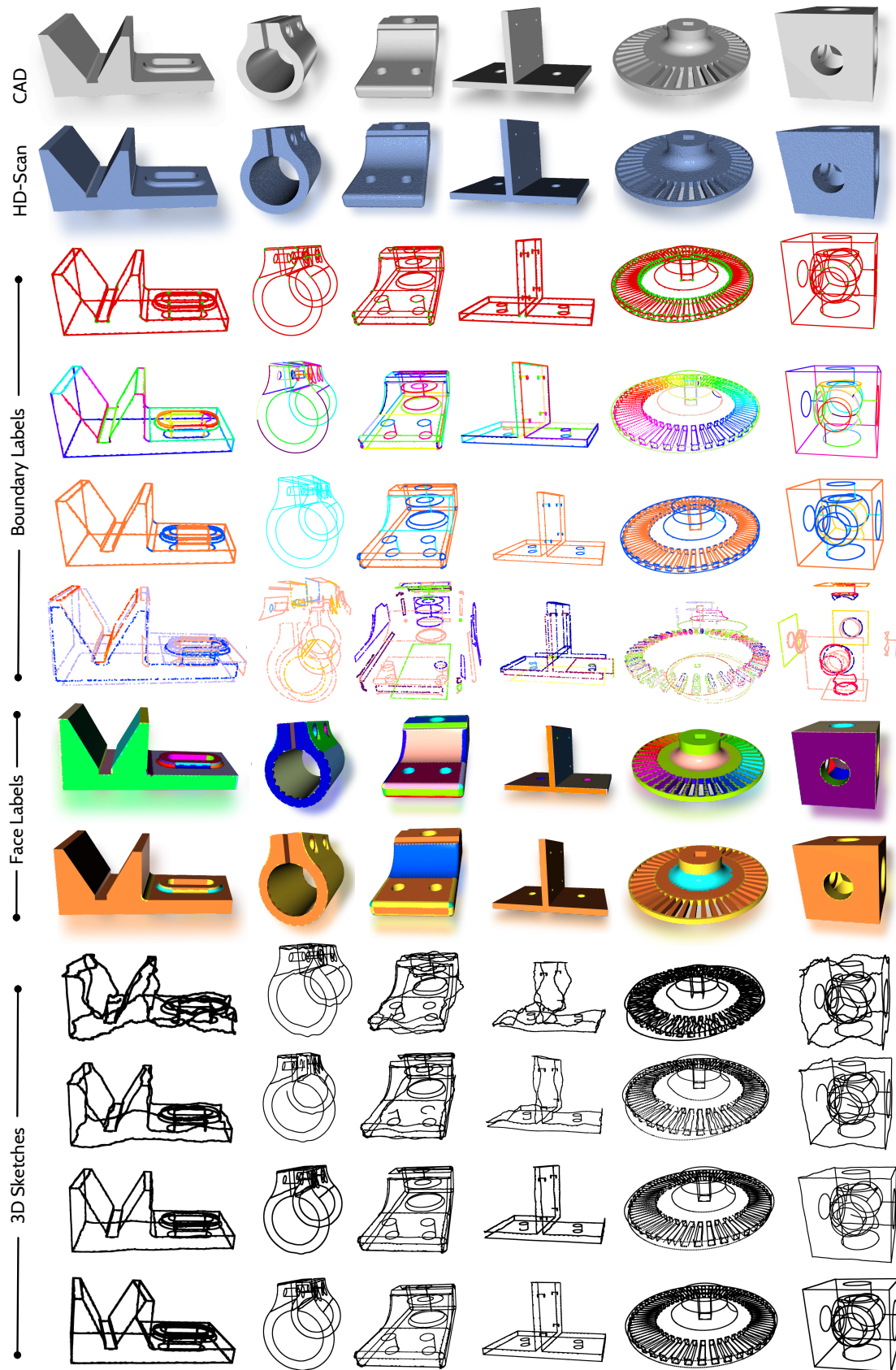


Figure 12. More visual examples of BRep Annotations for reverse engineering from 3D scans/sketches.

Chunk	Boundary		Junction		Chunk	Boundary		Junction	
	Recall	Precision	Recall	Precision		Recall	Precision	Recall	Precision
0	0.972	0.885	0.690	0.740	50	0.974	0.888	0.642	0.728
1	0.972	0.905	0.710	0.774	51	0.980	0.900	0.670	0.723
2	0.980	0.883	0.681	0.703	52	0.967	0.885	0.665	0.739
3	0.977	0.883	0.688	0.739	53	0.971	0.882	0.675	0.726
4	0.964	0.841	0.698	0.701	54	0.965	0.876	0.639	0.734
5	0.977	0.877	0.688	0.765	55	0.963	0.895	0.668	0.759
6	0.972	0.889	0.673	0.764	56	0.950	0.884	0.629	0.726
7	0.958	0.851	0.690	0.791	57	0.966	0.886	0.686	0.749
8	0.961	0.879	0.682	0.708	58	0.970	0.888	0.673	0.743
9	0.972	0.877	0.682	0.731	59	0.970	0.885	0.645	0.748
10	0.982	0.878	0.692	0.780	60	0.965	0.885	0.641	0.756
11	0.973	0.878	0.689	0.710	61	0.975	0.876	0.657	0.742
12	0.965	0.868	0.698	0.711	62	0.979	0.887	0.668	0.720
13	0.980	0.895	0.678	0.737	63	0.972	0.882	0.674	0.743
14	0.970	0.900	0.696	0.736	64	0.967	0.867	0.641	0.676
15	0.974	0.890	0.683	0.747	65	0.967	0.911	0.682	0.760
16	0.939	0.876	0.703	0.740	66	0.965	0.885	0.640	0.716
17	0.974	0.864	0.682	0.680	67	0.963	0.878	0.637	0.757
18	0.958	0.874	0.674	0.778	68	0.978	0.888	0.684	0.748
19	0.973	0.886	0.679	0.745	69	0.974	0.894	0.657	0.763
20	0.959	0.870	0.702	0.755	70	0.981	0.909	0.781	0.897
21	0.971	0.881	0.688	0.707	71	0.978	0.892	0.775	0.917
22	0.979	0.886	0.680	0.765	72	0.980	0.890	0.777	0.931
23	0.976	0.891	0.678	0.747	73	0.987	0.909	0.782	0.928
24	0.979	0.898	0.683	0.771	74	0.983	0.900	0.776	0.902
25	0.978	0.876	0.681	0.748	75	0.983	0.899	0.772	0.893
26	0.963	0.880	0.688	0.736	76	0.985	0.896	0.777	0.938
27	0.963	0.857	0.695	0.739	77	0.986	0.884	0.783	0.909
28	0.956	0.870	0.695	0.692	78	0.986	0.913	0.775	0.873
29	0.970	0.887	0.681	0.763	79	0.980	0.895	0.777	0.941
30	0.967	0.891	0.695	0.799	80	0.985	0.900	0.784	0.940
31	0.970	0.885	0.651	0.744	81	0.981	0.893	0.775	0.917
32	0.972	0.896	0.658	0.777	82	0.981	0.896	0.741	0.917
33	0.964	0.848	0.656	0.715	83	0.987	0.901	0.782	0.932
34	0.979	0.901	0.658	0.714	84	0.984	0.905	0.779	0.844
35	0.965	0.873	0.658	0.726	85	0.985	0.917	0.767	0.869
36	0.980	0.895	0.651	0.714	86	0.981	0.883	0.744	0.929
37	0.958	0.852	0.648	0.794	87	0.983	0.895	0.740	0.912
38	0.968	0.871	0.621	0.759	88	0.984	0.905	0.742	0.895
39	0.975	0.872	0.674	0.724	89	0.986	0.895	0.740	0.872
40	0.978	0.903	0.660	0.731	90	0.977	0.894	0.748	0.913
41	0.974	0.866	0.662	0.756	91	0.981	0.896	0.748	0.931
42	0.967	0.886	0.659	0.731	92	0.984	0.885	0.752	0.915
43	0.964	0.874	0.656	0.700	93	0.984	0.891	0.742	0.899
44	0.948	0.861	0.653	0.751	94	0.983	0.902	0.742	0.934
45	0.972	0.872	0.659	0.727	95	0.986	0.903	0.741	0.915
46	0.966	0.888	0.673	0.766	96	0.982	0.877	0.745	0.912
47	0.976	0.894	0.638	0.726	97	0.983	0.886	0.778	0.920
48	0.979	0.881	0.639	0.741	98	0.982	0.884	0.779	0.924
49	0.961	0.864	0.635	0.721	99	0.977	0.870	0.781	0.893

Table 4. Precision-recall measures for all 100 chunks of **A2Z** using our foundation model.