



# JanusCoder & PrismaCoder: Pushing the Boundaries of Open Multimodal Code Intelligence



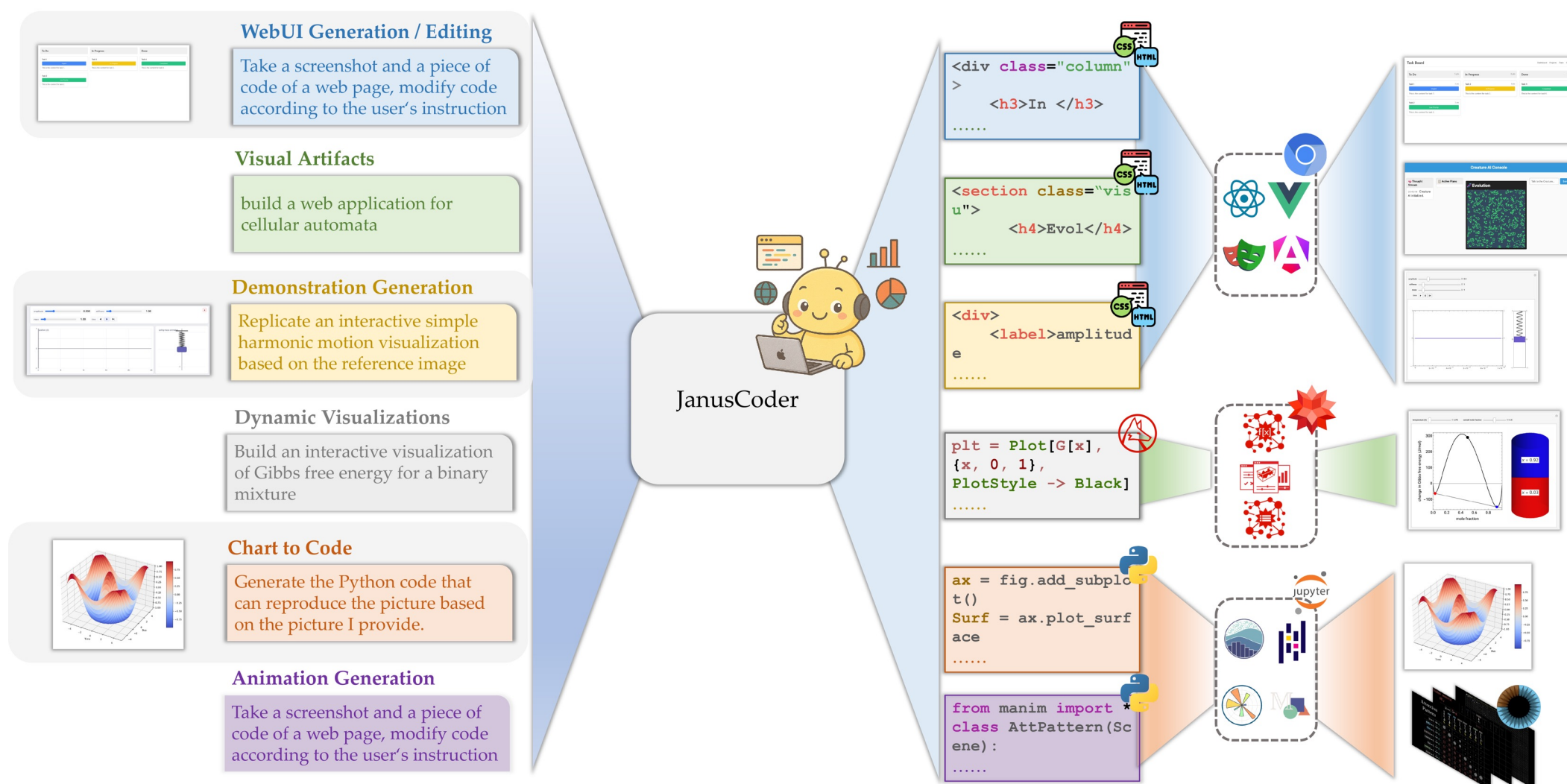
## Where We Started

## JanusCoder (ICLR 2026)

Code intelligence now reaches **the visual world programs produce**: charts, web UIs, scientific demos, animations.

Open work covered **one corner at a time**, and multimodal code data was scarce and unevenly curated.

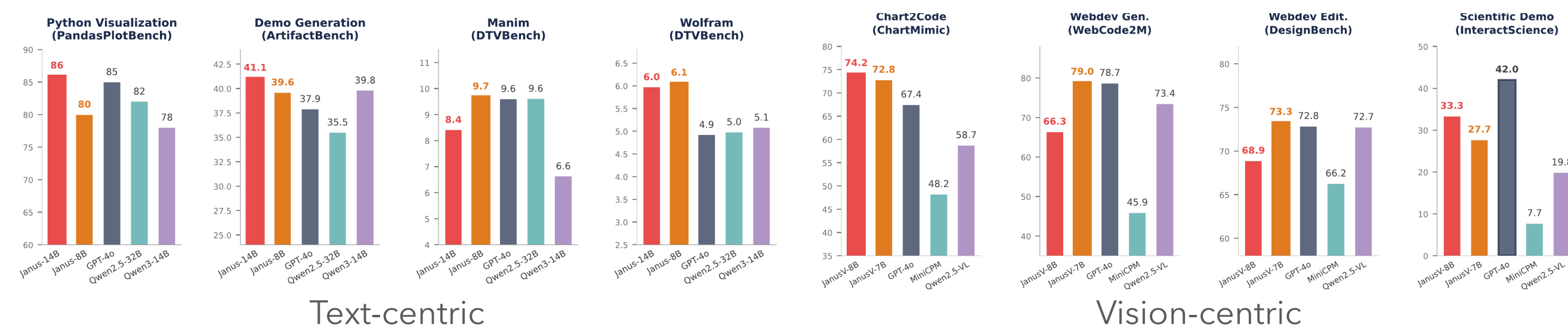
**JanusCoder: one unified visual-programmatic interface** for text- and vision-centric tasks, trained on JanusCode-800K.



One model, many formats: the unified visual-programmatic interface.

## How JanusCoder Performs

## Results



JanusCoder (8B/14B) and JanusCoderV (7B/8B) lead open models on plotting, artifacts, chart-to-code, and web generation. Pipeline, data, and checkpoints released.

## What JanusCoder leaves open

- Data is kept by one holistic score per sample, which says little about why a sample fails.
- Training is SFT only, with no signal that reads the render.
- Coverage stops at 800K samples over a handful of task families.

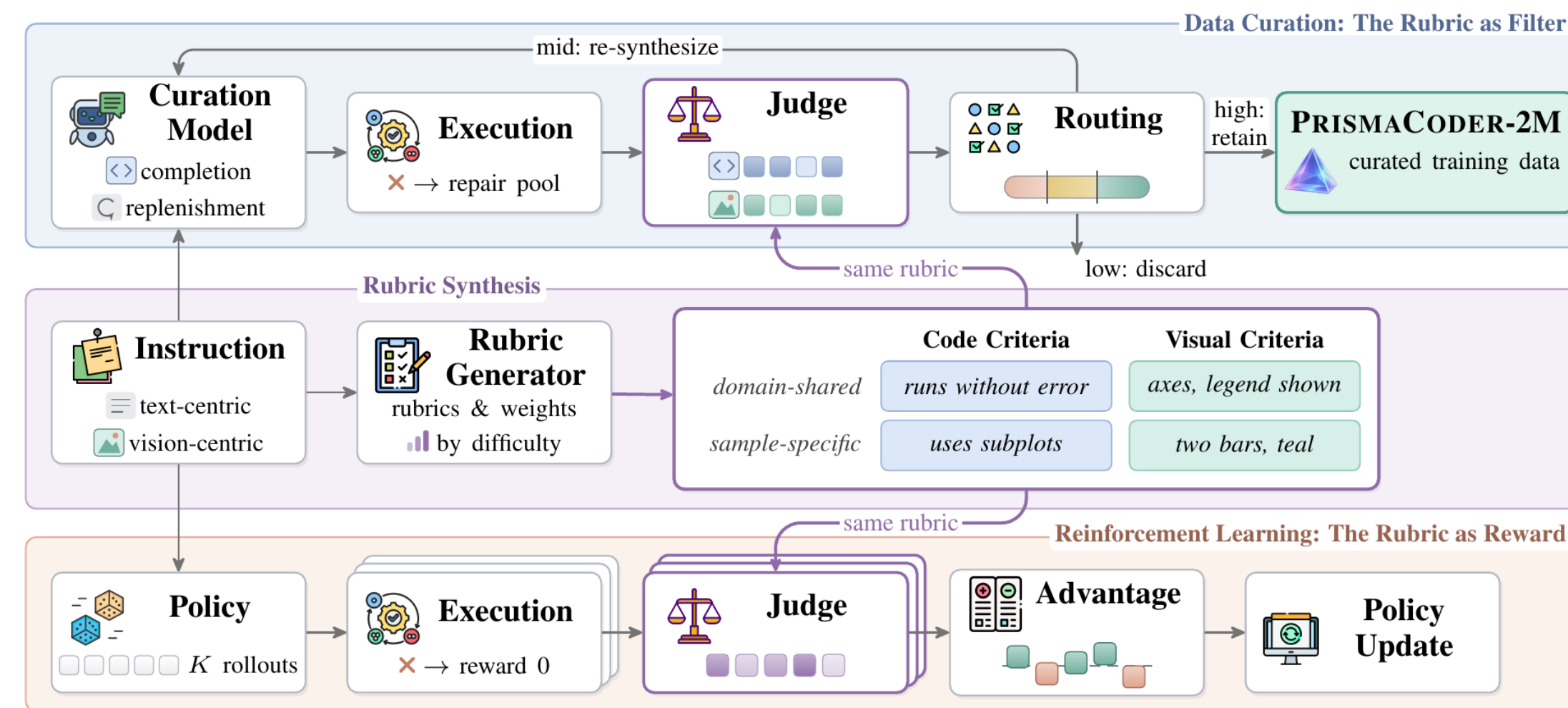
From JanusCoder to **PrismaCoder**, what changed

one score per sample → one rubric per instruction  
800K → 2M samples, 18 families, 10 languages

SFT only → rubric-rewarded RL  
7B to 14B → 9B and 35B-A3B

## How We Do That

## PrismaForge



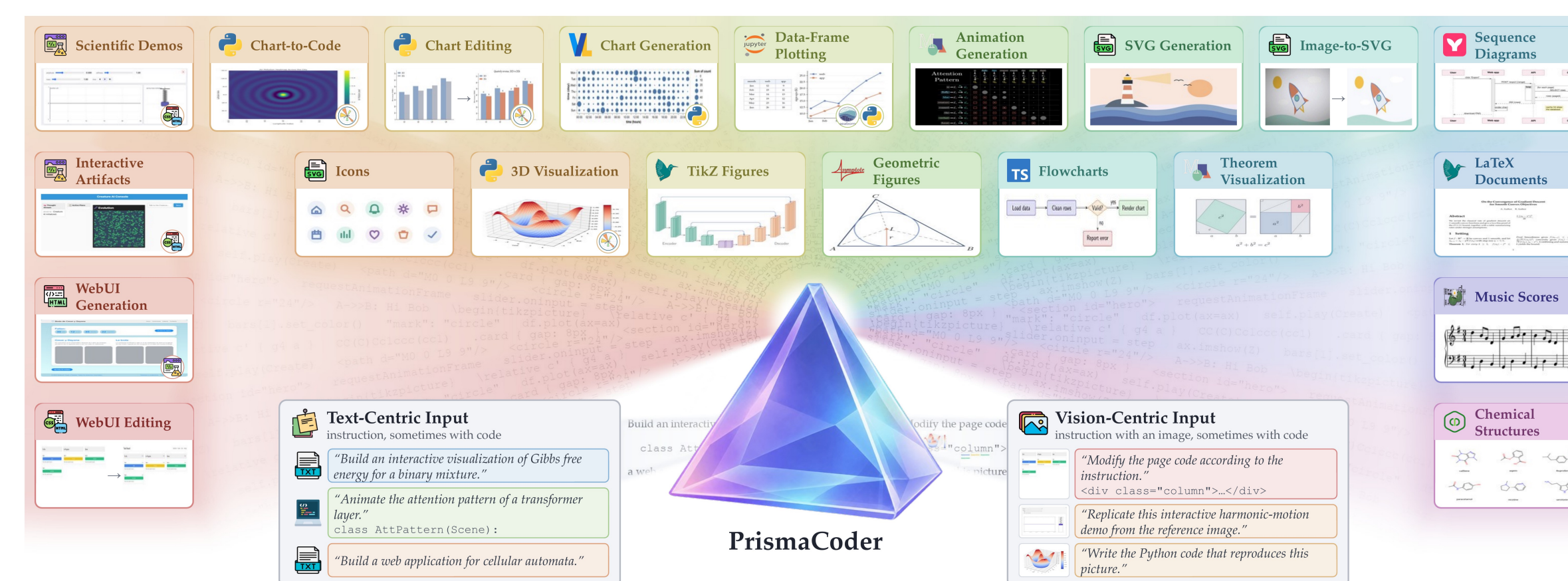
**Per-instruction rubric**  
shared + task-specific items, scored per modality

**Rubric as filter**  
discard, re-synthesize with feedback, or retain

**Rubric as reward**  
a dense signal for group-relative RL (GRPO)

## What We Release

## PrismaCoder-2M



**2M**

samples, text- and vision-centric

**18 / 10**

task families / languages

**59 : 41**

vision- to text-centric split (%)

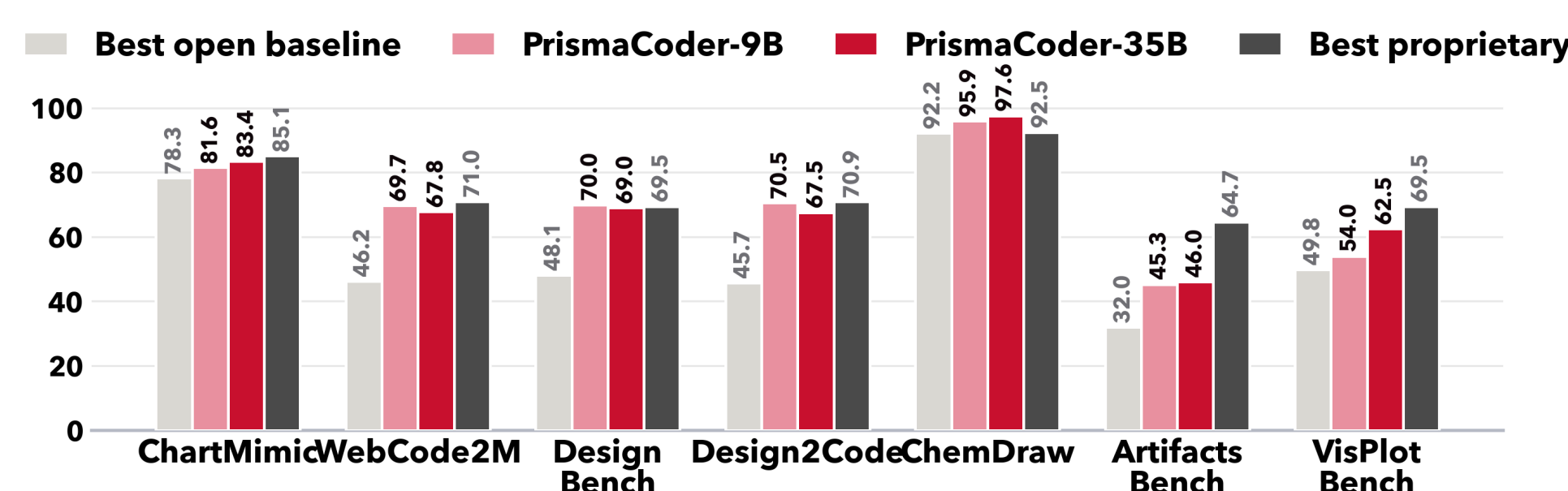
**13 of 15**

visual data types, most of any open work

The broadest and largest open dataset for multimodal code intelligence, with rubric labels on every sample, curated by the same rubric that later rewards RL.

## How PrismaCoder Performs

## Experiments



**Lead every open baseline** on nearly every benchmark, by a wide margin.

**Level with proprietary systems** on web interfaces; ahead on chemical-structure code.

**Rubric filtering beats direct scoring** at equal data size; **RL adds on every metric** on two other backbones.

