WORLD2DATA - ONE-PAGE TECHNICAL OVERVIEW

1. PROBLEM & CHALLENGE

Many AI and robotics teams cannot quickly convert real-world video into structured 3D data that is simulation-ready. Existing pipelines often require LiDAR, manual labeling, or brittle one-off scripts. This slows iteration and blocks reproducible evaluation.

2. TARGET AUDIENCE

Primary users:
- Robotics and embodied-AI teams needing fast world modeling.
- Simulation/digital-twin engineers who need OpenUSD-compatible scene assets.

Secondary users:
- Applied AI teams that need temporal object/state traces for downstream reasoning.

3. SOLUTION & CORE FEATURES

World2Data ingests 2D video and produces:
- Keyframes and camera trajectory.
- Dense 3D point cloud (MASt3R).
- Temporal scene graph with object/state metadata.
- OpenUSD export (`.usda`) for simulation workflows.
- Rerun recording (`.rrd`) for timeline playback and debugging.
- PLY export for standard point-cloud tools.

4. UNIQUE SELLING PROPOSITION (USP)

Single-video to OpenUSD with temporal scene context, confidence-aware reasoning hooks, and submission-ready visual outputs in one pipeline. No specialized capture hardware is required.

5. IMPLEMENTATION & TECHNOLOGY

- Geometry: MASt3R multi-view reconstruction.
- Semantics/Detection: YOLOv8 (+ optional SAM3 path when access is granted).
- Reasoning: Gemini-backed semantic/causal augmentation with graceful fallback.
- Output/Interop: OpenUSD (`usd-core`) + Rerun + JSON scene graph.
- Reliability: pytest suite with fast tests and overnight long-run mode.

6. RESULTS & IMPACT

Recent run artifacts demonstrate:
- ~737k reconstructed 3D points.
- 20 temporal camera frames.
- Structured scene graph and USD export generated end-to-end.
- Robust completion even under external API quota constraints (degraded mode).

If we had 24 more hours: finalize full SAM3-gated path and ship stronger multi-object temporal persistence in the demo video.