

Autonomous Inspection of Wind Turbine Interiors



Multimodal digital twins and vision-language agents for AI supported preventive maintenance
Presented by Jack Landers, MS in AI at Santa Clara University

AUTONOMOUS SYSTEMS | RENEWABLE ENERGY | HUMAN-IN-THE-LOOP AI

CENTRAL IDEA A digital twin constructed from autonomous drone inspection lets an offline vision-language agent evaluate component specific, multi-view evidence - while technicians retain final safety authority.

PROBLEM AND INSPECTION STRATEGY

Technicians work at height near high-voltage, hydraulic, and rotating equipment, then enter confined nacelle passages. The strategy is to move reconnaissance, documentation, and repeatable visual screening off-tower without transferring final safety authority to the model.

844 US occupational deaths from falls, slips, and trips in 2024
> 90m tower height before nacelle access
0.6m clearance in portions of the target nacelle

REGISTERED EVIDENCE STACK

RGB IMAGERY
Appearance, component state, and source frames for technician review.

LIDAR DEPTH
Measured geometry for the reconstructed surface and occlusion checks.

POSE AND TRAJECTORY
Camera viewpoints transformed into a common spatial coordinate frame.

TELEMETRY
Flight state aligned with video timestamps to synchronize observations.

TASK ALLOCATION STRATEGY

SHIFT TOWARD REMOTE SCREENING

Structural damage, corrosion, and leaks
Cable condition and bolt torque-mark alignment
Visible sensor readings and documentation

RESERVE FOR TECHNICIANS

Physical measurements and repairs
Safety-critical functional tests
Validation or override of every preliminary assessment

STRATEGY PRINCIPLES

Evidence before inference
Conclusions are limited to localized views of the relevant component.

Corroborate across views
Multiple observations and explicit provenance support technician review.

Route uncertainty to people
Unobservable tasks and uncertain grades move to technician or AR-assisted follow-up.

TECHNICAL METHOD

A single drone flight becomes a spatially registered evidence system that is queried task by task.

- Capture synchronized evidence**
Record RGB imagery, LiDAR, trajectory, and telemetry during a predetermined autonomous indoor flight.
- Align time and coordinate frames**
Match video timestamps to drone telemetry, estimate frame-level poses, and transform observations into one reference frame.
- Construct the spatial digital twin**
Apply Marching Cubes surface reconstruction to measured LiDAR depth, creating a 3D mesh while retaining links to source images and camera poses.
- Convert checklist tasks into evidence specifications**
Represent each maintenance item as a target component, observable condition, and task-specific prompt.
- Retrieve views by camera pose**
Select candidate frames whose viewpoints indicate that the queried component was observed.
- Reject geometrically occluded viewpoints**
Use mesh geometry to exclude frames where intervening surfaces block the target region.
- Localize and associate multi-view evidence**
Segment the target in selected images and associate observations with the 3D reconstruction using structure-from-motion geometry.
- Assess, cite, and route for review**
Pass localized views and the task specific prompts to the vision-language model that returns a preliminary grade, cited frames, and rationale for technician approval.

OBSERVABILITY GATE

Is the requested condition physically observable in the retrieved evidence?

YES Preliminary grade + cited frames + rationale

NO / UNCERTAIN Route to technician follow-up for AR assisted review

SYSTEM OUTPUTS

- Interactive 3D digital twin reconstructed from one drone flight
- Structured report for each evaluated maintenance task
- Preliminary grade with cited frames and supporting rationale
- Queue of incomplete or uncertain tasks for AR-assisted technician work

PRELIMINARY EVIDENCE

681 annual inspection tasks compiled as candidates for partial automation

PROJECTED FIELD-STUDY TARGETS

40% less power-generation downtime
80% less technician time inside the tower

Targets - not measured outcomes.

WHAT THE PROTOTYPE ESTABLISHED

- End-to-end capture, reconstruction, task query, and reporting pipeline
- Common frame registration of imagery, LiDAR, trajectory, and telemetry
- Component localized retrieval with repeated checklist conditioned queries
- Technician facing review of the model response and cited evidence

VALIDATION GAPS AND FAILURE MODES

- No controlled expert study of assessment accuracy yet
- Localization drift in repetitive component scenes
- Occlusion and low-light imagery
- Limited thermal, acoustic, and electrical interpretation
- Fleet-scale operational and economic impact remain unproven

NEXT FIELD STUDY

Test reliability across turbine models, seasons, environmental conditions, and maintenance teams. Quantify expert agreement, operational accuracy, and economic impact.

TAKEAWAY Our prototype organizes embodied sensing into traceable maintenance evidence: geometry filters what can be seen, the model evaluates only localized observations, and technicians remain the final safety authority.