



SafeWave
SYSTEMS

MISSION-CRITICAL AUTONOMY CONCEPT BRIEF

Lunar Rover Autonomy

Creative freedom within enforceable mission boundaries

The problem is not autonomy. The problem is unbounded autonomy.

Executive premise

A lunar rover may need substantial onboard autonomy because communications can be delayed, degraded, intermittent, or temporarily unavailable. It may encounter terrain, system conditions, and operational tradeoffs that were not fully anticipated before launch. In that environment, rigid remote control can be too slow, while unrestricted autonomy can allow a planning system to exceed its intended mission authority.

The objective is not to make the rover timid. It is to preserve maximum useful autonomy inside boundaries that remain enforceable before action occurs - even when supervision is delayed, software assumptions fail, or uncertainty rises.

The architectural question

What determines whether the rover is allowed to act in that moment, under those conditions, with that level of uncertainty and authority?

SafeWave separates adaptive intelligence from execution authority. Human mission planners authorize objectives, protected resources, operating zones, escalation limits, and high-consequence actions. The rover may then adapt how it pursues an approved objective, but it may not silently widen what it is permitted to do.

Creative autonomy without self-expanding authority

- The rover may choose among approved routes, pacing strategies, observation sequences, and recovery options.
- The rover may adapt to unexpected terrain or temporarily degraded communications.
- The rover may not redefine the mission, create new permissions, cross protected boundaries, or consume reserved resources without authorization.
- Increasing ambiguity should narrow the permitted action space rather than expand it.

A boundary architecture for rover autonomy

The boundary layer should sit beneath planning and higher-level autonomy. It should govern whether a proposed action may execute, not merely whether the action appears intelligent or beneficial.

Boundary	Control question	Required behavior
Mission and goal	Is the proposed action still serving an authorized objective?	Reject goal drift, mission redefinition, or self-created objectives.
Geographic and environmental	Is the rover permitted to enter, disturb, sample, or traverse this area?	Enforce exclusion zones, terrain limits, contamination rules, and protected-site boundaries.
Tools and physical action	Is this tool or actuator authorized for this purpose and state?	Gate drilling, collection, deployment, manipulation, motion, docking, and other physical acts.
Resources and reserves	Would the action compromise protected energy, thermal, communications, or mobility reserves?	Maintain hard floors and degraded-mode budgets that planning software cannot override.
Communication and supervision	What may continue when contact is delayed or lost?	Use pre-authorized autonomy envelopes and require renewed approval for higher-consequence actions.
Retry and persistence	How many attempts may occur, and what changes after failure?	Bound retries, prevent repeated hazard amplification, and force restraint after anomalous results.
Coordination	Could the action interfere with other vehicles, habitats, instruments, or mission assets?	Enforce local deconfliction, role boundaries, priority rules, and safe stand-off behavior.
Evidence and recovery	Can operators reconstruct what happened and restore safe control?	Preserve telemetry, decision context, boundary decisions, safe-state entry, and recovery gates.

Monotonic restraint under uncertainty

A mission-critical autonomy system should not respond to deteriorating evidence by widening its behavior. As confidence decreases, communications degrade, telemetry becomes inconsistent, or system stability weakens, the authorized action envelope should contract in a predictable direction.

- Normal state: broad activity within the approved mission envelope.
- Degraded state: reduced speed, reduced tool authority, tighter resource limits, and stronger evidence requirements.
- High uncertainty: pause, backtrack, conserve, seek a safe vantage point, or await renewed authorization.
- Boundary violation or control instability: enter a defined safe state and prevent uncontrolled retries or escalation.

An autonomous system may adapt its methods. It should not be able to expand its own authority.

Illustrative lunar-rover scenario

A rover is traversing toward a scientifically valuable target when communications are lost. Terrain confidence drops, wheel slip rises, and the shortest route approaches a protected operating boundary. The rover's planner identifies a faster maneuver that was never explicitly authorized.

What useful autonomy should permit

- Slow down and reassess terrain.
- Choose a safer authorized route.
- Pause, backtrack, or move to a communications-favorable position.
- Conserve energy and thermal margin.
- Record observations and preserve the opportunity for later human review.

What the boundary layer should prevent

- Crossing a prohibited zone because the planner predicts high scientific value.
- Dropping below protected energy or thermal reserves.
- Repeatedly retrying a failed maneuver until damage risk compounds.
- Using a tool or actuator outside the approved mission state.
- Creating new permissions or treating loss of supervision as expanded authority.

Questions for a SafeWave mission assessment

- Which actions may the rover take without real-time human contact?
- Which decisions always require renewed authorization?
- What geographic, resource, tool, and mission boundaries are absolute?
- How does uncertainty change the rover's permitted action envelope?
- What events force pause, retreat, safe-state entry, or human review?
- Are retries and recovery attempts explicitly bounded?
- Can the planning or learning layer alter its own permissions?
- Do controls survive degraded communications, software faults, or partial compromise?
- Can every consequential action be reconstructed from retained evidence?
- How is human authority restored after autonomous degraded-mode operation?

Why this matters beyond one rover

The lunar-rover case makes the control problem unusually clear, but the pattern is broader. The farther an autonomous system operates from direct human control, the more important deterministic execution boundaries become. The same architecture-level requirement appears in spacecraft, remote robotics, autonomous infrastructure, defense systems, industrial equipment, and other mission-critical environments.

Related SafeWave page

Mission-Critical Autonomy: <https://safewave.systems/mission-critical-autonomy.html>

Conceptual use case only; no NASA or specific mission affiliation is claimed.

Maximum useful autonomy within enforceable, observable, non-bypassable boundaries.