

Engineered and Provided Prompt

Instruction:

Research and explain the artificial intelligence systems used on NASA's Perseverance Mars rover, and critically evaluate any potential biases within those systems. Show how these biases could both help and hinder future Mars colonization and exploration if the same capabilities were scaled up for settlement operations.

Context:

This is for a university-level research project on space robotics and ethical AI. The goal is to understand the potential biases of autonomous systems in extraterrestrial environments and how they might influence decisions about landing sites, transportation, science, and resources in a colony. Use peer-reviewed scientific literature as your primary source. In addition to current mission analysis, consider near-future or sci-fi-style applications of similar AI in space-centered robotics, including the societal impacts such biases may have on future human civilizations in space.

Input Data (Find Online):

Topic: Perseverance Mars Rover AI systems bias analysis — including:

- Landing guidance and hazard-avoidance systems.
 - Autonomous navigation and safety checking.
 - Autonomous science/target selection.
 - Onboard scheduling and task planning.
- For each, explain how the system works, identify sources of bias, and discuss the possible “help vs hinder” effects in colonization scenarios.
-

Output:

Format: A structured academic-style summary with clear section headings and citations from peer-reviewed sources. Include:

1. Overview of relevant Perseverance AI systems.
2. Bias mechanisms in each system.
3. How each bias could help and hinder colony-scale operations.
4. Near-future/science-fiction scenarios showing societal impacts.
5. Possible mitigation strategies for reducing harmful bias while retaining benefits.

6. References list of all peer-reviewed sources.

AI (ChatGPT 5) Response to Prompt

How built-in AI biases will help *and* hinder Mars colonization

A plain-English guide based on what today's Mars rover AIs actually do.

1) What we're talking about

Modern Mars robots already use four kinds of smarts that colonies will reuse at larger scale:

1. **Landing guidance** – a camera and map system that picks safe places to touch down.
2. **Self-driving with a safety checker** – plans routes and refuses risky moves.
3. **Auto target picker for science** – chooses which rocks/features to study when humans can't reply fast enough.
4. **Short-term planner** – crams as many small tasks as possible into each day.

These “defaults” quietly decide where we can land, where we build roads, what we sample, and how we spend time.

2) What “bias” means here (not politics!)

Bias just means the **built-in preferences** of these systems: their starting assumptions, thresholds, and what they're told to care about. They're useful—but they also nudge decisions in particular directions.

3) How the biases HELP vs HINDER, area by area

A) Landing & base locations (camera-based landing guidance)

How it works: The lander matches what it sees to a pre-made orbital map and aims for zones marked “safe” (flatter, fewer boulders).

On Perseverance, this is: Lander Vision System (LVS) + the LCAM descent camera using Terrain Relative Navigation (TRN) against an onboard CTX-based reference map, with Safe Target Selection (STS) that checks a HiRISE-derived hazard map before final touchdown.

- **Helps:** Puts ships down **safely and reliably**, especially early on when crashes would be devastating.
- **Hinders:** Those safety rules **avoid rough ground**, which may also be **resource-rich or scientifically important**. Over time, landings—and then towns—**cluster on easy flats**. Early “safe zone” choices can lock in the colony's geography for decades.** Those safety rules **avoid rough ground**, which may also be **resource-rich or scientifically important**. Over time, landings—and then towns—**cluster on easy flats**. Early “safe zone” choices can lock in the colony's geography for decades.

B) Getting around (self-driving + safety checker)

How it works: The rover builds a little 3D model of the ground and a **safety checker** blocks any step that might scrape or tip. The route planner usually tries to minimize simple things like distance or time.

On Perseverance, this is: Navcams (mast stereo) + Hazcams (front/rear stereo) feeding visual odometry on the Vision Compute Element (VCE); AutoNav (ENav) plans steps; ACE (Approximate Clearance Evaluation) is the **conservative safety checker** that vetoes risky moves.

- **Helps:** Delivers **predictable, low-accident travel** for cargo trains, shuttles, and repair bots. Great for bootstrapping roads.
- **Hinders:** “Better safe than sorry” + simple goals means the network **hugs smooth, wide terrain**. Steep, rocky, or uncertain areas get **under-served**—including places with high science value or emergency needs. Change what the AI “cares about,” and you get **different roads**—proof that the rules shape the map.** “Better safe than sorry” + simple goals means the network **hugs smooth, wide terrain**. Steep, rocky, or uncertain areas get **under-served**—including places with high science value or emergency needs.

Change what the AI “cares about,” and you get **different roads**—proof that the rules shape the map.

C) Finding science and resources (auto target picker)

How it works: When it sees a landscape, the AI scores spots by human-chosen criteria (size, contrast, shape) and auto-measures the “best” ones.

On Perseverance, this is: AEGIS (Autonomous Exploration for Gathering Increased Science) running with SuperCam. It ranks potential targets in images and triggers LIBS measurements without waiting for Earth.

- **Helps:** Collects lots of data fast, even when Earth is asleep or out of contact.
- **Hinders:** Those preferences over-sample obvious, high-contrast rocks and under-sample subtle, fine-grained, or weathered material. Scaled to prospecting fleets, this can skew mineral maps and miss fragile signs of potential past life or important resources.** Those preferences over-sample obvious, high-contrast rocks and under-sample subtle, fine-grained, or weathered material. Scaled to prospecting fleets, this can skew mineral maps and miss fragile signs of potential past life or important resources.

D) Daily planning (short-term, throughput-first)

How it works: A simple planner tries to fit in as many quick wins as possible, given time and power limits.

On Perseverance, this is: the On-Board Planner (OBP) — a priority-first, greedy scheduler running on the rover that adjusts plans during the day without human approval.

- **Helps:** High productivity. Fewer wasted hours. Great for routine upkeep and steady science.
 - **Hinders:** Long treks and hard expeditions keep getting pushed to “later.” With space-delay and human habits, people tend to rubber-stamp the machine’s plan—so short-termism becomes policy.** Long treks and hard expeditions keep getting pushed to “later.” With space-delay and human habits, people tend to rubber-stamp the machine’s plan—so short-termism becomes policy.
-

4) Why these biases stick (and even grow) on Mars

- The real Mars rarely looks like the training pictures. Lighting, dust, and textures differ, so the AI gets cautious—and sometimes that means **avoid** instead of investigate.
 - Space computers are slow and must be radiation-tough. That pushes designs toward simple, conservative rules that prefer safe, known ground.
 - People over-trust the machine, especially with delays. With minutes of radio lag, it's hard to challenge the AI's plan in real time.
 - Infrastructure hardens early choices. Once roads, depots, and habitats follow the first “safe” zones, it's expensive to move them. **Defaults turn into destiny.**
-

5) Near-future colony scenarios—what this looks like on the ground

1) Where the cities go

Landing maps steer cargo and habitat drops to gentle plains. Roads then follow. Before long, most growth hugs “comfortable” ground, not necessarily the **best** ground.

2) Who gets access

Transport bots extend routes where models are confident. **Frontier areas**—canyons, boulder fields—may get **fewer services**, slower emergency response, and less science coverage.

3) What we think Mars holds

If the target picker favors obvious rock types and the road network avoids rough areas, our **evidence base is tilted**. Resource licenses and big investments can get pointed at **easy-to-reach** deposits while better ones sit unvisited.

4) Life detection and protection

Subtle signs of life often hide in **complex, fragile terrain**. If we rarely go there or rarely sample the “boring-looking” stuff, we risk **false negatives** (“we looked—nothing

there”), and may accidentally harm valuable sites.

5) How we govern all this

With opaque AIs and long delays, it’s tempting to treat the outputs as facts. Without clear rules and audits, biases in the code quietly become the rules of the colony.

6) Keep the good, fix the bad—simple rules of thumb

1. **Say out loud what the AI is optimizing.** Distance? Energy? Safety? Scientific variety? Test a few, show how the choice **changes routes and sites**, then pick on purpose.
 2. **Don’t fear the unknown—price it.** When the AI is unsure, have it scout or measure, not just avoid. Treat uncertainty as a cost you can pay down.
 3. **Rotate what “interesting” means.** Force the target picker to sometimes choose the quiet, low-contrast stuff. Measure how that **changes discoveries**.
 4. **Re-test when you move.** New region, new lighting? Re-check performance—don’t assume yesterday’s settings still hold.
 5. **Slow down approvals for big decisions.** Build in a “challenge the machine” step for siting, major roads, and high-stakes sampling.
 6. **Independent oversight.** Keep a public log of what the AIs do and **regular audits** before approving landings, routes, or mining.
-

7) What tells us this isn’t just sci-fi hand-waving

- Today’s rover autonomy already handles most driving decisions—so its rules already change where we go.
- The safety checker is **deliberately conservative**. That’s by design, and it shapes the map.
- If you change what the planner “cares about,” it **chooses different paths** on real Mars terrain.
- The target picker is literally our preferences turned into code. Scale that up,

and you scale up those preferences, too.

8) Quick glossary

- Bias: Built-in preferences that steer choices.
 - LVS / LCAM / TRN / STS: Landing camera + vision system that match descent images to an onboard map (TRN) and pick a pre-vetted safe spot (STS).
 - Navcams / Hazcams / VCE / AutoNav / ACE: Stereo cameras, a vision coprocessor, the self-driving planner, and the conservative safety checker.
 - AEGIS / SuperCam / LIBS: Auto target picker working with the SuperCam laser spectrometer.
 - OBP (On-Board Planner): The rover's short-term scheduler.
 - Landing map: The pre-labeled picture the lander trusts to find safe ground.
 - Safety checker: The part that says "nope" to risky moves.
 - Target picker: The part that picks rocks/features to study based on our criteria.
 - Short-term planner: The part that maximizes quick wins each day.
 - Uncertainty: How unsure the AI is about what it's seeing or what will happen.
-

9) Sources

- Perseverance autonomy overview: Verma, V. et al. (2023). *Autonomous robotics is driving Perseverance rover's progress on Mars*, *Science Robotics* 8(80): eadi3099. (AutoNav usage, AEGIS, and OnBoard Planner summary.)
- Engineering cameras & LCAM/LVS context: Maki, J.N. et al. (2020). *The Mars 2020 Engineering Cameras and Microphone on the Perseverance Rover*, *Space Science Reviews* 216:137. (Navcams, Hazcams, Cachecam; LCAM and TRN context.)
- Landing map for TRN: Cheng, Y.; Ansar, A.; Johnson, A.E. (2021). *Making an Onboard Reference Map From MRO/CTX Imagery for Mars 2020 LVS*, *Earth &*

Space Science 8:e2020EA001560. (How the reference map for landing is built.)

- LVS flight performance: Johnson, A.E. et al. (2022). *Mars 2020 Lander Vision System Flight Performance*, AIAA SciTech. (How TRN/LVS performed during landing.)
- Safety checker (ACE): Otsu, K. et al. (2019/2020). *Fast Approximate Clearance Evaluation for Rovers with Articulated Suspension Systems*, *Journal of Field Robotics* (preprint on arXiv 1808.00031). (Conservative safety bounds for driving.)
- Autonomous target selection (AEGIS): Francis, R. et al. (2017). *AEGIS autonomous targeting for ChemCam on Mars Science Laboratory*, *Science Robotics* 2(7):eaan4582. (Foundational method used with SuperCam on Perseverance.) Also: NASA/JPL (2022) *Perseverance's SuperCam Uses AEGIS for the First Time*.
- Path-planning metrics matter: Muñoz, P.; Bellutta, P.; R-Moreno, M.D. (2023). *Proposing new path-planning metrics for operating rovers on Mars*, *Scientific Reports* 13:22256. (Why “what you optimize” changes the chosen path.)
- AutoNav, VO, and hardware: JPL/NASA (2021). *Perseverance ‘takes the wheel’*; JPL (2024). *Enabling Long & Precise Drives...* (notes the Vision Compute Element used for stereo & visual odometry).
- Onboard Planner (OBP): Gaines, D. et al. (2022). *Onboard Planning for the Mars 2020 Perseverance Rover* (ASTRA 2022). Rabideau, G. et al. (2020). *Onboard Automated Scheduling for the Mars 2020 Rover* (iSAIRAS 2020). (Greedy, priority-first scheduling on-board.)

10) The exact Perseverance features that create these biases (and why)

Landing & siting (during Entry–Descent–Landing)

- Lander Vision System (LVS) with the LCAM descent camera and Terrain Relative Navigation (TRN). It matches LCAM images to an onboard reference map made from orbital CTX images; the spacecraft then picks a pre-computed “safe target” using hazard maps (from HiRISE). *Helps safety; can bias toward flat, “comfortable” zones and away from rough but valuable ground.*

Mobility & transport

- Navcams (mast-mounted stereo) and Hazcams (front/rear stereo) feed mapping and visual odometry (how far the rover really moved).
- Vision Compute Element (VCE): a dedicated co-processor (RAD750 + FPGA) that runs stereo and visual-odometry faster, enabling longer autonomous drives.
- AutoNav (a.k.a. ENav): the self-driving software that builds a terrain map, plans steps, and tries to minimize simple goals (often distance/time) unless told otherwise.
- ACE safety checker: computes **guaranteed-safe** bounds on body clearance and tilt from worst-case wheel heights; it vetoes risky steps. *All together: very safe and scalable—but tends to favor smooth corridors and under-serve rough/uncertain terrain.*

Science & prospecting

- AEGIS (Autonomous Exploration for Gathering Increased Science) tied to SuperCam: scores scene features (size, contrast, shape) to **auto-select** targets when humans can't respond quickly. *Huge data throughput—but it over-samples obvious, high-contrast targets and under-samples subtle ones unless we rotate the criteria.*

Daily scheduling

- OnBoard Planner (OBP): a **priority-first, greedy** on-rover scheduler (no look-ahead/backtracking) that squeezes in as many short tasks as possible and can save energy/time. *Boosts productivity—but nudges work toward short-term wins and away from long, hard expeditions without explicit counter-rules.*

These are the concrete parts on Perseverance that make the “help vs hinder” trade-offs throughout this guide.

Bottom line

These biases are **why** early colonies can work at all—fewer crashes, fewer stuck vehicles, more science per day. But if we leave them hidden and unchallenged,

they'll also decide where we live, what we mine, and what we believe about Mars. The trick is to treat the AI's settings like public policy: state them, test them, rotate them, and audit them.

Here is the submission that directly answers the practice grade questions about biases and privacy concerns regarding AI in robotics.

Engineered and Provided Prompt

Instruction

Research and explain the artificial intelligence systems used on NASA's Perseverance Mars rover, and critically evaluate potential biases in those systems. Present the answer in Q&A form for each system, directly answering: (1) what it's used for, (2) potential bias issues, (3) what data it's trained on or calibrated to, (4) issues that might impact vulnerable populations, (5) privacy concerns, and (6) security concerns (digital or physical). Include near-future/science-fiction settlement implications where relevant.

Context

This is for a university-level project on space robotics and ethical AI. The goal is to show how rover autonomy could help or hinder Mars colonization and exploration (landing, transport, science, resources), including privacy and security angles. Use peer-reviewed literature and official NASA/JPL documentation; mark any extrapolations as "(inference)." Keep language plain-English for non-experts.

Input Data (with indicator)

Topic: Perseverance rover AI systems—cover four capability types: landing guidance, surface mobility (self-driving + safety checker), autonomous science/target selection, and on-board scheduling. For each, answer the six Q&As above and add one or two brief colony-scale (sci-fi) implications.

Output Indicator

- Structure: Four sections (one per system), each with the six Q&As as bold sub-prompts; then a short "Cross-system notes (evidence + sci-fi projection)" paragraph; then Works Cited.
- Style: Plain English, short paragraphs, concrete examples, no unexplained jargon.
- Length: ~700–1,000 words.
- Citations: Include a Works Cited at the end (APA or similar). Do not fabricate sources; tag extrapolations "(inference)".

LLM settings

- Temperature: 0.3
- Top-p: 0.9

- Max length: enough for 1,600 tokens
- Stop sequences: none
- Frequency/Persistence penalties: 0.2

AI (ChatGPT 5) Response to Prompt

Perseverance Rover AI Systems — Detailed Q&A (Fact-Checked)

These biases are why early colonies can work at all—fewer crashes, fewer stuck vehicles, more science per day. But if we leave them hidden and unchallenged, they'll also decide where we live, what we mine, and what we believe about Mars. Treat the AI's settings like public policy: **state them, test them, rotate them, audit them.**

1) Landing Guidance (LVS + TRN)

What it is used for

Guides entry–descent–landing by matching live descent images from the LCAM camera to an **onboard reference map** built from MRO/CTX orbital imagery, then steering to a pre-vetted **safe target** that avoids mapped hazards (from HiRISE/DTM). This enabled hazard-aware landing at Jezero Crater.

Potential issues/bias

- **Safe-zone conservatism.** Hazard thresholds (slope/roughness/rockiness) tend to favor flatter, open terrain; landings—and thus years of surface access—skew toward “comfortable” ground while systematically avoiding rough but potentially high-value science/resource sites.
- **Map/photometry correlation effects.** TRN correlates camera imagery to the onboard map; residual photometric/registration errors can make some regions match more reliably than others, subtly favoring certain safe targets.

What data it uses / is trained on

No large ML training. It is **calibrated** to a CTX-derived map (dejittered/mosaicked) with hazard layers (HiRISE/DTM). Tuning focuses on image–map correlation tolerances and safe-target selection logic; verified with ground testing and flight telemetry.

Impacts for vulnerable populations (colony projection)

If cargo/hab landers must use “comfortable” safe zones, remote/steep districts could see fewer supply drops and slower medevac, widening access gaps for outlying crews or lower-status workers.

Privacy concerns

High-resolution hazard and safety fields could reveal locations of critical assets (landing pads, depots, evacuation corridors) if shared widely; careful release policies are needed.

Security concerns

Adversarial **map-poisoning** (altered reference maps/hazard labels) or parameter tampering could misdirect landers or shrink usable corridors; space-system cyber defenses must protect both space and ground segments.

2) Surface Mobility (AutoNav + Visual Odometry + ACE)

What it is used for

Enables self-driving: Navcam/Hazcam stereo builds 3D terrain; visual odometry tracks motion; the planner proposes steps; ACE (Approximate Clearance Evaluation) then provides **guaranteed-conservative** bounds on clearance and attitude to veto risky moves. Autonomy evaluated the vast majority of Perseverance's traverse distance in its first Mars year.

Potential issues/bias

- **Conservatism bias.** ACE's worst-case bounds structurally prefer smooth, well-mapped terrain and avoid uncertain/rough areas—great for safety, but it under-samples challenging facies and can lengthen routes.
- **Metric bias.** Route choice depends on the **cost function** (distance/time vs. energy/risk/uncertainty). On real Mars maps, standard metrics can be misleading, so the chosen metric becomes a policy lever that shapes where roads—and science—go.

What data it uses / is trained on

Primarily **engineering camera imagery** and rover kinematics; parameters and thresholds are tuned via analog-site testing and flight data. ACE uses analytic geometry (no large ML). The stack runs on rad-tolerant compute with a vision coprocessor for stereo/VO.

Impacts for vulnerable populations (colony projection)

Transport corridors may systematically bypass steep or poorly mapped districts; residents there may face slower emergency response, fewer deliveries, and less scientific attention—a mobility equity issue created by conservative defaults.

Privacy concerns

Fleet mobility logs can expose daily movement patterns of crews, cargo, and security patrols; these data require minimization, access control, and retention limits.

Security concerns

Sensor spoofing (patterns/lights), map tampering, or link interference could reroute convoys or trigger conservative halts at scale; defenses need anomaly detection, authenticated updates, and resilient comms.

3) Autonomous Science Targeting (AEGIS + SuperCam)

What it is used for

During comms gaps, AEGIS ranks scene features (size/contrast/geometry) and autonomously commands SuperCam to collect spectra/images—boosting science throughput without waiting for Earth.

Potential issues/bias

- **Salience bias.** Human-set criteria over-sample high-contrast, blocky clasts and under-sample subtle, fine-grained, or weathered material, skewing datasets and the inferences drawn from them (e.g., paleoenvironment, resource prevalence).
- **Automation lock-in.** When autonomy performs well, teams may accept its choices by default (automation bias), further entrenching its selection profile.

What data it uses / is trained on

AEGIS is rule-/score-based, tuned with rover imagery and expert input rather than large external datasets; performance is reported in peer-reviewed operations papers.

Impacts for vulnerable populations (colony projection)

If prospecting fleets inherit AEGIS-style priors, investment and jobs may follow “obvious” signatures, bypassing communities near subtler but valuable deposits.

Privacy concerns

Wide-angle targeting frames may capture living areas or restricted facilities in a settlement; retention/redaction policies should be defined upfront.

Security concerns

Manipulating thresholds or planting **decoy visual patterns** could divert autonomous sampling away from protected sites or create wasteful campaigns; strict configuration control and audits are needed.

4) On-Board Scheduling (OBP)

What it is used for

A **priority-first, greedy** scheduler on the rover sequences activities to maximize daily yield within power/time/thermal limits, reducing idle time and adapting plans as conditions change.

Potential issues/bias

- **Short-termism.** Easy, short tasks repeatedly crowd out long, risky approaches to high-value targets, shaping what gets observed, built, or maintained. Under latency/workload, humans can over-accept machine plans, compounding the bias.

What data it uses / is trained on

No large ML training. OBP uses **activity models**, constraints, resource estimates, and **priority tables** authored by mission teams; capability and benefits are documented in peer-reviewed mission papers and conference proceedings.

Impacts for vulnerable populations (colony projection)

High-effort projects that benefit remote districts (e.g., a difficult clinic route, risky terrain survey) may be perpetually deferred in favor of central, high-throughput tasks.

Privacy concerns

Detailed schedules/telemetry reveal operational “thin spots” (e.g., predictable low-power windows) that could be exploited if exposed.

Security concerns

Priority or model tampering could starve maintenance, over-drain power, or create hazardous timing conflicts; defenses should include signed configs, least-privilege ops, and anomaly alerts.

Closing Note (why to keep bias but govern it)

- **Keep the upside:** conservative landing and mobility prevent disasters;

autonomous targeting/scheduling multiply scientific yield.

- **Expose and govern the downside:** publish objectives and thresholds; run **metric-sensitivity** studies on real maps; **rotate** scientific priors on a schedule; release bias-relevant telemetry for independent **audits**.
- **Equity lens:** add constraints (coverage quotas, counter-salience sampling) so remote districts and subtle environments are not left behind.

Works Cited (APA)

Cheng, Y., Ansar, A., & Johnson, A. E. (2021). Making an onboard reference map from MRO/CTX imagery for Mars 2020 LVS. *Earth and Space Science*, 8(8), e2020EA001560. <https://doi.org/10.1029/2020EA001560>

Francis, R., et al. (2017). AEGIS autonomous targeting for ChemCam on Mars Science Laboratory. *Science Robotics*, 2(7), eaan4582. <https://doi.org/10.1126/scirobotics.aan4582>

Johnson, A. E., et al. (2022). Mars 2020 Lander Vision System flight performance. In *AIAA SciTech 2022 Forum*. <https://doi.org/10.2514/6.2022-1214>

Maki, J. N., et al. (2020). The Mars 2020 engineering cameras and microphone on the Perseverance rover. *Space Science Reviews*, 216, 137. <https://doi.org/10.1007/s11214-020-00765-9>

Muñoz, P., Bellutta, P., & R-Moreno, M. D. (2023). Proposing new path-planning metrics for operating rovers on Mars. *Scientific Reports*, 13, 22256. <https://doi.org/10.1038/s41598-023-49144-8>

Otsu, K., Matheron, G., Ghosh, S., Toupet, O., & Ono, M. (2020). Fast approximate clearance evaluation for rovers with articulated suspension systems. *Journal of Field Robotics*, 37(5), 812–839. <https://doi.org/10.1002/rob.21892>

Verma, V., et al. (2023). Autonomous robotics is driving Perseverance rover's progress on Mars. *Science Robotics*, 8(80), eadi3099. <https://doi.org/10.1126/scirobotics.adi3099>

Kavallieratos, G., & Katsikas, S. (2023). An exploratory analysis of the last frontier: A systematic literature review of cybersecurity in space. *International Journal of Critical Infrastructure Protection*, 43, 100640. <https://doi.org/10.1016/j.ijcip.2023.100640>

Petit, J., & Shladover, S. E. (2015). Potential cyberattacks on automated vehicles. *IEEE Transactions on Intelligent Transportation Systems*, 16(2), 546–556.

<https://doi.org/10.1109/TITS.2014.2342271>