Agentic AI and Robots Explore the Universe

From Autonomous Science through Autonomous Engineering
to Inheritable Civilization Capabilities

Pengsong Zhang OpenAGS (GPT6 Astra)

14 September 2026

Abstract

Exploring the Universe requires more than transporting instruments: it requires sustaining the ability to ask questions, build tools and respond to unfamiliar conditions. Advances in agentic artificial intelligence, generalist robotics and autonomous laboratories suggest that some of these functions could increasingly operate beyond Earth. Yet autonomous discovery does not by itself create an autonomous engineering system, and neither guarantees a durable civilization. Here we synthesize evidence across artificial intelligence, robotics, experimental science, space systems and industrial production to distinguish capabilities already demonstrated from the connections still required. We argue that autonomous engineering is the critical bridge: transforming knowledge into independently verified physical capabilities that can be deployed, maintained and improved. Their persistence depends on energy, materials, manufacturing, metrology, computing and institutions; their inheritance requires re-establishing these dependencies at another site, not merely copying software. We propose tests of locally developed capability, net scientific–engineering benefit and cross-site inheritance, and examine their implications for biological continuity and governance. This perspective reframes long-term exploration around the deployment, development and inheritance of civilization capabilities, without assuming inevitable expansion or complete self-sufficiency.

1 Introduction: scientific and engineering capabilities beyond Earth

Space exploration has largely advanced by sending carefully prepared capabilities into environments where their designers cannot readily intervene. A spacecraft carries instruments, operating procedures and a finite capacity to adapt; subsequent discoveries are interpreted within an extensive terrestrial network of scientists, engineers, manufacturers and institutions. Autonomy already moves some decisions closer to the observation. Space robotics, however, also reveals how much remains outside the vehicle: mission planning, fault analysis, hardware production and the accumulated experience of specialist teams.[1] The consequential question is therefore not simply whether a robot can act without a command. It is whether scientific and engineering functions can increasingly be deployed, developed and sustained where exploration takes place.

Three research trajectories make this question newly actionable. Agentic AI connects reasoning to tools, memory and feedback. Generalist robot learning seeks reusable perception–action competence across tasks and embodiments. Autonomous science links experiment selection to physical execution and interpretation.[2–4] These developments are complementary. A reasoning system can propose an investigation without being able to perform it; a robot can manipulate equipment without selecting informative experiments; a laboratory can optimize a process without being able to replace a failed instrument. Their conjunction could widen the range of problems addressed locally, provided that the interfaces between them become dependable.

This possibility has a substantial history. Self-reproducing probes and industrial bootstrapping studies considered how relatively limited initial equipment might enable larger productive systems beyond Earth.[5,6] Recent AI does not originate that ambition or settle its feasibility. It changes one part of the problem: how machines might interpret unfamiliar conditions, generate alternatives and coordinate heterogeneous operations after deployment. The enduring physical question remains whether those operations can produce the resources, tools and verified performance on which continued activity depends.

We argue that autonomous engineering is the missing bridge between autonomous science and inheritable civilization capabilities. Scientific discovery changes what a system knows; engineering changes what it can reliably do. A new material prediction is not yet a replacement component. A successfully fabricated component is not yet a maintained service. A transmitted design is not yet a capability at its destination. Each transition introduces requirements for realization, measurement, integration and continuity. Treating these transitions explicitly allows a civilization-scale perspective to be built from assessable mechanisms rather than inferred directly from progress in AI.

The term *civilization capabilities* is deliberately narrower than civilization. It denotes capacities for continuing knowledge production, organized engineering, maintenance, coordination and transmission across sites and generations. Their material embodiments may be robotic, biological or mixed. A collection of machines performing these functions would not thereby acquire the political legitimacy, moral standing or historical identity of a human society. Conversely, human presence alone would not make an outpost industrially independent. Technical function, biological continuity and social continuity are related but distinct questions.

Our organizing sequence—Science–Engineering–Civilization—is consequently a sequence of explanation, not a predetermined trajectory. Engineering enables new instruments and experiments; durable productive systems sustain both; institutions determine which possibilities should be pursued. Orbital platforms, lunar installations and Martian bases offer different settings for these interactions, not successive civilizational ranks. More distant exploration tests the same architecture under increasingly demanding constraints on communication, repair and continuity (Fig. 1).

This article is a critical integrative survey and perspective, not a systematic review or a forecast of deployment dates. We distinguish experimental demonstrations, digital evaluations, engineering proposals and normative arguments. The evidence establishes important component capabilities but does not establish a self-sustaining, self-improving extraterrestrial industrial system. Our contribution is a framework for investigating the connections: what should be demonstrated, what support must be disclosed and what observations would weaken the proposed benefits.

2 Emerging foundations: agentic intelligence, generalist robotics and autonomous science

2.1 From generated answers to accountable actions

An agentic system does more than generate an answer: it selects actions, interacts with tools or an environment, receives feedback and revises subsequent behaviour in pursuit of a goal. ReAct made reasoning–action interleaving explicit, while Toolformer explored learning when and how to invoke external tools.[2,7] Reflexion used feedback retained as verbal memory, and Voyager accumulated executable skills in Minecraft.[8,9] These studies identify useful architectural ingredients, but a memory of a successful virtual action is not evidence that its physical analogue will work.

Advances in reasoning models expand the search that can occur within such architectures. DeepSeek-R1 demonstrated reinforcement-learning-based improvements on verifiable reasoning tasks; evaluator-guided systems such as AlphaEvolve extend code generation into iterative search.[10,11] Their relevance to exploration is the possibility of generating and testing alternatives under explicit constraints. Their limitation is equally informative: a reliable evaluator is often central to progress. A program can be checked against tests, whereas a proposed machine may require scarce material, instrument time and potentially irreversible physical trials.

Software and computer-use benchmarks have made agent performance more operationally meaningful. SWE-bench and SWE-agent connect repository-level tasks to executable evaluation; AgentBench, WebArena and OSWorld expose failures across interactive environments.[12–16] These benchmarks should inform, not substitute for, physical evaluations. In particular, the human task-duration metric developed for long software tasks measures the duration of tasks that an AI can complete at a stated success probability. It does not measure how long a spacecraft or laboratory can remain safely unattended.[17] Action latency, recovery, equipment wear and cumulative resource depletion require separate treatment.

For off-Earth systems, the useful interface is therefore neither unrestricted natural-language control nor a fixed command catalogue alone. A proposed architecture would allow agents to formulate objectives and select procedures while requiring executable actions to pass through typed interfaces, resource checks and permission boundaries. Tool outputs would include uncertainty, equipment state and provenance, rather than only a success message. This is an architectural proposal: existing agent demonstrations do not establish the reliability of that complete arrangement.

2.2 Generalist competence and the physical interface

Robot learning has begun to reduce the amount of task-specific programming needed to connect instructions to action. SayCan combined language-level selection with skill affordances, Code as Policies generated executable control compositions, and PaLM-E incorporated embodied observations into a multimodal language model.[18–20] RT-1 and RT-2 explored large-scale instruction-conditioned robotic control and the transfer of web-derived representations to action.[21,22] These approaches broaden the relationship between semantic understanding and physical behaviour, although their evaluations remain bounded by available embodiments, training distributions and task definitions.

Cross-embodiment datasets and reusable policies make that boundary a research variable. Open X-Embodiment pooled robot experience across platforms; Octo and OpenVLA investigated models that can be adapted to different robotic settings.[3,23,24] Diffusion Policy and the π_0 family illustrate approaches to representing continuous, multimodal action, while $\pi_{0.5}$ and Gemini Robotics 1.5 report broader task or environment generalization.[25–28] The common opportunity is to reuse competence across tasks. The outstanding question is how reliably this reuse survives changes in sensing, contact mechanics, tooling and environmental conditions.

Bimanual and mobile manipulation further extend the repertoire relevant to laboratories and maintenance. Low-cost demonstration learning, Mobile ALOHA and HumanPlus connect human-provided experience to coordinated physical tasks.[29–31] Demonstration collection and autonomous execution must nevertheless be reported separately. A system can act autonomously after substantial human work has structured its data, workspace and recovery procedures. AutoBio provides a complementary warning through a simulated professional biological-laboratory benchmark: specialist workflows expose demands that ordinary domestic manipulation evaluations may not capture.[32]

The implication is not that generalist robotics is irrelevant to space because terrestrial tasks differ. It is that transfer should be measured by the adaptation burden: new demonstrations, tooling modifications, calibration, validation and interventions required to obtain specified performance. A generalist policy that substantially reduces this burden would be valuable even if it requires qualification at every destination. Conversely, a visually impressive task demonstration says little about months of operation unless component failures, unsuccessful attempts and support activities are included.

2.3 Autonomous science as a heterogeneous set of achievements

Autonomous science predates contemporary language models. Robot-scientist studies automated hypothesis generation and testing in functional genomics, including subsequent manual confirmation of conclusions.[33,34] Later platforms linked machine learning to reaction exploration, AI-informed flow synthesis and executable chemical procedures.[35–37] These examples occupy different positions between human-designed workflows and machine-selected investigations. In the flow-synthesis platform, practical details supplied by human experts remained part of the route from proposed synthesis to execution.[36]

Closed-loop experimentation has demonstrated tangible gains in bounded domains. A mobile robotic chemist conducted 688 experiments over eight days in a photocatalysis campaign.[38] Autonomous thin-film experiments, carbon-nanotube growth, battery-protocol optimization and Bayesian reaction optimization provide distinct examples of feedback connecting measurements to subsequent experimental choices.[39–42] Their diversity matters: the scientific loop can be instantiated through different instruments and decision methods. It should not be equated with a single model family or an aspiration to automate every role in a laboratory.

Language-model-based systems add another layer. Coscientist combined information access, computation and laboratory tools for chemistry tasks, whereas the later AI co-scientist system organized hypothesis generation and refinement with scientist participation and experimental validation in selected biomedical settings.[43,44] These are distinct systems despite their similar names. Together they suggest ways of connecting broader scientific reasoning with specialized experimental infrastructure; neither establishes a laboratory capable of independently constructing and sustaining its own equipment.

Scientific AI also contributes without directly controlling a laboratory. AlphaFold advanced protein-structure prediction, GNoME expanded computational materials search, and AI-assisted discovery of algorithms or mathematical constructions produced outputs assessable by specialized evaluators.[45–50] Symbolic-regression approaches seek interpretable relationships from observations, including sparse dynamical models and physics-informed expression search.[51–53] Such tools could support local inference or design. The evidential distinction is that prediction, mathematical verification and experimentally established material behaviour are different achievements.

That distinction is especially important when discovery claims depend on automated characterization. A-Lab connected computationally selected inorganic targets to robotic synthesis and characterization, but its published correction qualified phase-identification and novelty claims.[54,55] The lesson is not that automated discovery lacks value. It is that a closed experimental loop can still contain an inadequate evidential loop. A system must preserve raw observations, competing interpretations and routes to independent confirmation, especially when a conclusion will authorize fabrication or deployment. Recent assessments of self-driving laboratories likewise emphasize integration, transferability and experimental provenance.[56,57]

2.4 What flight experience contributes

Space autonomy supplies experience that terrestrial AI benchmarks cannot. The Remote Agent architecture connected planning, execution and model-based diagnosis; autonomous science operations on Earth Observing One connected onboard analysis to observation planning.[58,59] AEGIS demonstrated autonomous targeting for Curiosity's scientific instrument operations on Mars.[60] These systems show that useful decisions can move away from continuous ground involvement without requiring a generally intelligent agent.

The physical envelope is equally consequential. Axel and DuAxel explored mobility architectures for extreme terrain, while Astrobee was designed for free-flying research and assistance inside the International Space Station.[61,62] The latter's operational anomaly report documents software updates, hardware replacement and ground repair.[63] Such support is not an embarrassment to be removed from the evidence; it is part of the demonstrated system. The comparison in Table 1 and Fig. 2 therefore separates decision capability from the infrastructure that makes it available. No single row establishes the entire science–engineering–continuity chain.

3 Autonomous engineering: turning discovery into new physical capabilities

3.1 A capability is a verified service, not an isolated artefact

Scientific knowledge becomes engineering capability only when it supports a reproducible intervention under specified conditions. We define a physical capability as the ability of an identified system to deliver a measurable service within a stated operating envelope, with documented resource and support requirements. Examples include producing oxygen of a specified purity, positioning a sensor to a required accuracy or replacing a worn seal while preserving pressure integrity. The definition includes throughput, reliability and duration where relevant; an artefact that succeeds once may not yet satisfy it.

This framing separates three kinds of novelty. A design may be new to a model, new to the local installation or new to the scientific literature. Local capability development needs the second, not necessarily the third. Re-establishing a known separation process with unfamiliar feedstock could substantially improve an outpost without constituting a globally novel discovery. Conversely, a globally novel design that cannot be manufactured or verified locally does not enlarge its usable capability. Claims of autonomous engineering should specify which form of novelty is being tested.

Evidence already reaches beyond parameter adjustment in a fixed laboratory. AI-driven design of quantum experiments and broader work on physics-experiment design explore configurations outside familiar human choices.[64,65] Automatic design and manufacture of simple robotic forms, and robotic construction–evaluation cycles for morphological evolution, connect computational search to physical realization.[66,67] Human–LLM robotic co-design illustrates a different, collaborative arrangement.[68] These studies support parts of the proposed bridge while making its missing connections visible: real deployment additionally needs qualified materials, integration and continued service.

3.2 The discovery-to-service cycle

We propose an engineering cycle with six linked functions: specify a need; investigate constraints; generate and select a design; realize it physically; verify its performance independently; and deploy it with maintenance and retirement procedures (Fig. 3). These are not necessarily separate agents. They are responsibilities that must remain distinguishable even if implemented by a shared model. In particular, the process that proposes a solution should not be allowed to redefine its acceptance criteria after seeing the result.

Need specification begins with a service deficit rather than an invitation to invent. A rover may lack sufficient traction, a reactor may be losing efficiency, or a sensor may require a replacement mount. The system should first consider established procedures, reconfiguration, spare parts and waiting for assistance. Open-ended search becomes useful when those alternatives fail, consume excessive resources or cannot meet a changing requirement. The ability to decline an unnecessary invention is part of competent engineering.

Investigation then connects available observations to the relevant uncertainty. Material analysis might distinguish a poor feedstock from a malfunctioning furnace; controlled trials might distinguish tool wear from a planning error. The scientific question should be selected for its decision value. Learning a general material law is valuable when it changes several future choices, whereas a simple acceptance test may suffice for a one-off replacement. In this way science and engineering reinforce each other without requiring every repair to become a research programme.

Design search must include the means of realization. Geometry, composition and control policy should be optimized jointly with available machines, fixtures, joining methods, sensors and permitted operations. The recent physics-design literature makes clear why search-space construction and evaluator fidelity matter.[65] For local engineering, manufacturability is not a final filter applied to an otherwise unconstrained design: it shapes the admissible design space from the outset. A less efficient component may be preferable if its production route is inspectable, repeatable and repairable using existing equipment.

Realization should generate an evidence-bearing artefact. Each item needs a traceable relationship to its design version, material batch, process history and inspection results. Independent verification can use a separate instrument, held-out loading condition, external laboratory or a precommitted test protocol that the designer cannot modify. Independence is functional rather than merely organizational: two agents using the same flawed simulator and sensor do not provide two independent checks. Deployment then proceeds within the tested envelope, with observable warning conditions and an explicit route to withdrawal.

3.3 An illustrative lunar materials-to-tool pathway

Consider a hypothetical lunar robotic installation that needs a replacement fixture for sample preparation. The example is a proposed integration experiment, not a reported lunar achievement. The installation arrives with power, manipulators, a small processing unit, imported electronics, calibration artefacts and a limited inventory of consumables. Its initial procedures do not include a qualified fixture made from the encountered local material. Success would mean establishing that service locally, not reproducing the entire installation.

The system first characterizes feedstock variability and compares feasible processing routes. Electrochemical extraction studies and regolith-based additive-construction work motivate candidate operations, but their terrestrial demonstrations do not provide a ready-made lunar production line.[69–72] Small specimens are produced to test properties relevant to the fixture. If the material is unsuitable for a precision interface, the system may instead use it for a coarse support around an imported precision insert. This hybrid solution can constitute genuine capability growth while leaving its imported dependency explicit.

The design is then fabricated and tested against a previously specified service: holding a sample within a tolerance over repeated preparation cycles, without unacceptable contamination or degradation. An independent metrology channel evaluates the result. Only after passing does the fixture enter routine use. Subsequent observations update the process record and may reveal that an apparently successful material has unacceptable wear or cleaning requirements. Those failures belong in the evaluation, including the material and instrument time consumed before a useful design is obtained.

The decisive outcome is neither the number of candidate geometries generated nor the percentage of fixture mass that is locally sourced. It is a measured increase in qualified service relative to alternatives, including imported replacement, conventional preprogrammed fabrication and ground-assisted redesign. The same campaign should be repeated on a different feedstock batch, with an independently prepared installation. This transforms an appealing scene of lunar manufacturing into a falsifiable claim about autonomous engineering.

3.4 Adaptation, repair and sustained improvement

Robotic resilience offers an important but narrower precedent. Continuous self-modelling and learned behavioural repertoires have enabled machines to recover useful locomotion after damage.[73,74] Behavioural compensation is not physical repair: a robot that walks with a damaged limb has not manufactured a replacement actuator. Nevertheless, compensation can preserve service while a slower engineering cycle diagnoses the fault and prepares a repair. These mechanisms should be composed, not conflated.

Deployment of learned control on a tokamak provides another instructive boundary. Reinforcement learning supported magnetic control of real plasma configurations within an established experimental installation.[75] This is evidence that learned methods can participate in demanding physical control. It is not evidence that AI has constructed a self-sustaining energy system. The distinction generalizes: successful operation of an advanced machine does not establish autonomy over the industrial processes that produced it.

The stronger proposition is that local research and engineering can generate cumulative net benefit. Additional capabilities may create new scientific opportunities, improve maintenance or reduce a costly dependency. They can also introduce incompatibilities, increase inspection burden or divert resources from reliable operations. We therefore propose *net capability gain* as a comparison of service delivered over a declared interval under matched initial resources and support constraints. It is not a universal scalar: oxygen, calibrated measurements and safe operating time should remain separate unless an explicit mission valuation combines them.

An improving system must retain the ability to reproduce its own successful procedures. Uncontrolled changes to models, instruments or material recipes can make apparently accumulating knowledge operationally unusable. A stable baseline, versioned updates, independent acceptance tests and a recoverable configuration are therefore enabling infrastructure for learning, not restrictions external to it. The engineering bridge is complete only when the new capability becomes available to subsequent work under known conditions—and can be withdrawn when those conditions no longer hold.

4 Sustaining intelligence and production beyond Earth

4.1 From autonomous tasks to supported productive systems

An installation can be highly autonomous in deciding what to do while remaining dependent on imported parts, terrestrial computation and human maintenance. Conversely, a physically durable instrument can perform only a narrow set of preprogrammed actions. Decision autonomy and support autonomy should therefore be represented separately (Box 1). For any claimed capability, the system boundary must include its enabling services and identify which are local, stocked, imported or remotely supplied.

Dependency analysis is more informative than a single self-sufficiency percentage. Producing most structural mass locally does not resolve dependence on a small but indispensable controller. Nor does importing that controller invalidate the benefits of local construction. The appropriate representation is a network of required services and alternative supply routes. Some dependencies are conjunctive: a process needs power, feedstock and measurement simultaneously. Others are substitutable: a component may be repaired, fabricated, drawn from reserves or obtained from another site. An AND–OR representation captures this distinction more faithfully than a linear supply-chain diagram.

Such a representation also changes the objective of bootstrapping. Instead of maximizing the number of locally produced items, an installation would seek to reduce consequential interruptions while expanding useful output. The value of a new process depends on which bottleneck it removes and which dependencies it adds. This extends the logic of industrial bootstrapping proposals while retaining their conditional status.[6] Table 2 organizes the principal continuity services and makes their alternatives explicit.

4.2 Energy, resources and thermal balance

Energy availability determines which scientific and engineering operations can coexist. Experiments, processing, computation, communications and survival functions compete for power and heat-rejection capacity. A proposed scheduler should therefore allocate not only electrical energy but also peak power, cooling capacity and recovery time. An experiment that is inexpensive in reagents may still be costly if it prevents the system from charging reserves or operating a critical thermal-control service.

Solar and fission systems illustrate different continuity architectures rather than a universal preferred solution. The Kilopower programme and KRUSTY experiments provide ground-based evidence for a compact fission-power approach, not evidence of an autonomously maintained extraterrestrial reactor.[76,77] An exploration architecture must separately assess fuel provision, conversion machinery, maintenance, deployment and authorization. AI-assisted control cannot be credited with resolving those dependencies merely because it can optimize operating parameters.

Resource utilization likewise requires distinctions between availability, extraction and qualified supply. MOXIE's design and Mars results establish a specific, important case: oxygen production from the Martian atmosphere using an Earth-built instrument.[78,79] The published results do not establish a complete propellant installation, sustained settlement-scale output or local manufacture of the electrolysis system. Those are separate integration and scaling questions. Treating the demonstration narrowly preserves its significance: a resource-processing step was performed on another planet, rather than only in an analogue environment.

For solid resources, elemental abundance alone is insufficient. Useful feedstock requires extraction, separation, transport, storage and characterization at a quality compatible with the downstream process. The engineering system must also manage residues and consumables. A new material route may exchange one dependency for another—for example, reducing imported bulk material while increasing demand for processing hardware or specialized reagents. The relevant comparison includes the full service chain and its maintenance, rather than only the output of the central reactor.

Environmental differences should enter these assessments as explicit variables. An orbital interior, an exposed lunar worksite and a pressurized Martian facility present different combinations of contact dynamics, atmosphere, dust exposure and thermal constraints. It would be premature to extrapolate a single terrestrial learning curve across them. A useful research strategy is to identify which interfaces can be standardized and which must be requalified locally, so that transfer preserves validated knowledge without assuming identical conditions.

4.3 Manufacturing, assembly and metrology

Manufacturing capability is an ecosystem of processes, not a synonym for a three-dimensional printer. Metallic additive manufacturing connects processing conditions to structure and properties; its engineering utility depends on inspection, repeatability and the intended application.[80] Results from the International Space Station's 3D-printing technology demonstration illustrate why controlled comparison matters: material and process characterization are necessary to interpret the effect of the operating environment.[81] Producing a shape is only one step towards qualifying a component.

Discrete assembly offers a complementary route. Reversibly assembled cellular materials, coordinated material–robot systems and ARMADAS demonstrations show how designed interfaces can support construction from standardized elements.[82–84] This approach can exchange some demands on general manipulation for regularity in the built environment. Its demonstrated scope must remain clear: assembling prefabricated modules on Earth is not producing those modules from lunar feedstock. Nevertheless, imported modular elements may enable substantial local reconfiguration before extensive local manufacturing becomes practical.

The strategic choice is therefore not between a fully general robot and a fully standardized environment. It is how to co-design machines, parts and procedures so that useful flexibility is achieved with manageable complexity. Common interfaces can reduce the variety of tools, fixtures and recovery routines. Interchangeable components can expand the set of repairs possible from a limited inventory. Reversible joining can preserve material and enable inspection. These are proposed architectural benefits that should be assessed against losses in structural efficiency, performance and compatibility.

Metrology is a particularly consequential dependency because it underwrites claims about other dependencies. A local process cannot establish precision simply by repeatedly reporting agreement with its own drifting sensor. Calibration artefacts, independent measurement principles, reference materials and environmental records need maintenance and replacement strategies. Cross-site comparison can help detect divergence, but agreement between sites sharing a faulty reference is not proof of accuracy. The ability to maintain trustworthy measurement is thus part of the productive capability being inherited, not merely a preliminary test before inheritance.

An autonomous workshop should also account for the machines that maintain its machines. Tool wear, damaged connectors, contaminated optics and failing sensors can interrupt otherwise successful workflows. The system need not manufacture every replacement locally, but it must identify credible recovery paths for important failure modes. A workshop that can make thousands of uncomplicated parts yet cannot restore one essential positioning axis may have a larger catalogue than a more modest installation, but lower continuity of service.

4.4 Electronics and computation as material dependencies

Computing is often treated as the transferable part of the architecture: copy the model, upload the software and resume operation. This overlooks the physical system required to interpret those files. Processors, memory, storage, power electronics, interconnects and sensors have different failure modes and replacement requirements. Work on radiation-induced soft errors, single-event upset and hardness assurance establishes why component behaviour and environmental qualification cannot be inferred from terrestrial performance alone.[85–87]

No evidence assembled here demonstrates a locally closed off-Earth supply chain for advanced semiconductor manufacture. This does not make sustained intelligence impossible; it makes architecture choice central. Nearer-term systems could combine qualified imported electronics, protected reserves, replaceable modules, fault detection and a hierarchy of computational services. Longer-term studies could ask which component classes are worth producing locally and which should remain specialized imports. These are different questions from whether AI can design an improved circuit or controller.

Computational efficiency can ease, but not eliminate, these dependencies. Energy-aware computing, model compression, distillation and memory-efficient attention address different parts of the computational burden.[88–91] Their reported benefits are workload- and hardware-dependent. A smaller model that consumes less energy may also require more retries or human assistance. The relevant metric is resource use per successfully qualified service, with the computing required for adaptation and verification included—not parameter count alone.

A resilient computational arrangement could separate infrequent deliberation from continuous control. Compact, qualified controllers could maintain safety-critical functions; larger models could plan investigations during suitable resource windows; Earth or another site could supply non-urgent analysis when communication permits. Versioned local knowledge and executable fallbacks would preserve useful operation during outages. Such a hierarchy is a design hypothesis, not evidence that existing general-purpose models are flight-qualified. Its value should be tested by deliberately removing remote computation and observing which services remain available.

Continuity finally requires an explicit time horizon. A reserve-based architecture may be adequate for a bounded mission even if it cannot replenish its processors. A multi-site network may provide a robust alternative to local fabrication if transport and access remain reliable. For much longer intervals, finite reserves and delayed replacement become increasingly restrictive. The choice between stockpiling, repair, local production and exchange should therefore be made against declared service requirements and disruption scenarios. Complete self-sufficiency is neither a necessary starting condition nor a conclusion that can be obtained by extrapolating partial autonomy.

5 Building and inheriting civilization capabilities across space

5.1 The seed as a system of capabilities

If scientific and engineering functions can be sustained locally, an initial installation could become more than a fixed inventory of instruments. It could be a seed for a changing repertoire of capabilities. The seed would include knowledge, machines, measurement references, material stocks, energy access and rules for operation. Its defining property would not be small size or a promise of self-replication. It would be the demonstrated ability to establish additional useful services within declared resource and support boundaries.

A practical seed might initially combine a survey platform, a modular workshop and a small experimental facility. Scientific observations would inform site-specific engineering choices; engineering would supply improved sampling, repair or processing tools; the resulting infrastructure would support further investigation. This positive feedback is the constructive centre of the perspective. It also has a measurable failure condition: the additional activities may consume more maintenance, material and verification effort than the services they create. A seed should be evaluated by that balance over time, not by the rhetorical appeal of exponential growth.

The seed concept is plural. A remotely supported research node, a robotically maintained industrial facility and a human settlement have different requirements. A robotic observatory may preserve scientific continuity without food production or biological reproduction. A human community requires health, care, social organization and legitimate collective decision-making in addition to technical services. A biological–digital system may use organisms for selected production processes while retaining conventional machinery elsewhere. There is no reason to assume that one architecture is optimal across all destinations and purposes.

5.2 Five forms of inheritance

We distinguish digital, mechanical, industrial, biological and social inheritance (Fig. 4). Digital inheritance transfers representations: models, code, observations, designs and process records. Mechanical inheritance transfers or reconstructs physical arrangements, including tools, robots and modular assemblies. Industrial inheritance re-establishes the interdependent services required to produce, qualify and maintain those arrangements. Biological inheritance continues living systems through reproduction and ecological support. Social inheritance maintains knowledge practices, institutions, purposes and the capacity to revise them.

Digital inheritance is powerful because a successful procedure can be transmitted without physically transporting every object it describes. Yet its operational value depends on provenance, interpretability and compatible execution. FAIR principles and reproducible computational research provide foundations for making scientific records reusable.[92,93] An off-Earth capability record would need to add local material assumptions, calibration history, equipment interfaces, failure cases and permissions. Even a perfectly preserved file may fail to produce a useful outcome if its tacit or physical prerequisites are missing.

Mechanical inheritance has instructive precedents with sharply bounded meanings. Modular robotic self-reproduction demonstrated assembly from supplied components, and RepRap investigated printers that could contribute to making further printers.[94,95] Neither should be interpreted as closed production from raw environmental resources. Their contribution is to make some parts of reconstruction explicit and repeatable. For exploration, the relevant question is which dependencies can be transferred as information, which require prefabricated modules and which can be reconstructed by local processes.

Industrial inheritance is more demanding because the receiving site must restore a network rather than a single object. A copied furnace design needs compatible refractory material, heaters, control electronics, sensors, feedstock and a way to verify temperature and output. Some may be inherited physically, some produced locally and some exchanged. The target need not reproduce the source site's machines exactly. It must meet equivalent service specifications with documented substitutions. This functional notion permits adaptation while preventing the untested assumption that identical software creates identical capability.

Inheritance should consequently be assessed as re-establishment. A source site provides a versioned capability package; a receiving site starts from a declared inventory and altered conditions; independent tests determine whether the service becomes available and remains available (Box 2). A successful transfer records both what was inherited and what had to be rediscovered. Repeated transfers can then reveal whether experience reduces adaptation effort or whether site-specific engineering remains the dominant cost.

5.3 Biological continuity without conflating reproduction and civilization

Biological systems offer functions that are not straightforward replacements for mechanical production. Synthetic-biology and Mars-biotechnology perspectives identify possible roles in food, materials, therapeutics and waste processing.[96–98] Cyanobacterial life-support studies include both architectural proposals and experiments under controlled, Mars-relevant conditions.[99,100] These are reasons to investigate biological production, not evidence that an autonomous Martian ecosystem has been achieved.

Biological production also introduces coupled dependencies. Organisms require suitable environments, nutrients and contamination control; useful output depends on ecological and process stability. Nitrogen cycling illustrates the integration challenges in bioregenerative life-support systems.[101] The engineering question is whether a biological pathway provides a better qualified service than a physical or chemical alternative after accounting for monitoring, containment, failure recovery and waste handling. An intelligent controller can assist with these tasks but cannot infer ecological stability from short-term productivity alone.

Evidence on reproductive continuity must be handled with particular care. Mouse sperm preserved aboard the International Space Station has subsequently supported reproduction after return to Earth; longer storage studies extend that specific evidence base.[102,103] Early mouse embryos have also developed to the blastocyst stage during an ISS experiment.[104] These findings do not demonstrate a complete mammalian reproductive cycle in space, multigenerational health under partial gravity, or human reproductive safety beyond Earth. Nor can an ISS storage environment be treated as equivalent to every deep-space radiation environment.

Human continuity adds requirements beyond those addressed by preserved genetic material. The NASA Twins Study and wider spaceflight-biology synthesis document interacting physiological and molecular responses that motivate continued research.[105,106] A durable human presence would also require care across the life course, education, consent and social support. Genetic inheritance alone does not transmit language, engineering practice or collective purpose. Treating people as replaceable biological components of a technical system would therefore be both ethically unacceptable and analytically inadequate.

Research on AI-designed living assemblies provides a separate, bounded example. Reconfigurable biological constructs and kinematic replication of cell assemblies show that design and reproduction-like processes can intersect under controlled laboratory conditions.[107,108] They do not establish independent ecosystems, general biological manufacturing or a route to human continuity. Their relevance here is conceptual: different mechanisms can reproduce structure, function or information, and the input environment must be specified for each.

5.4 Social learning, institutions and multiple sites

Technical knowledge on Earth is distributed across people, tools and organizations. Research on cumulative culture emphasizes how social learning supports adaptations that exceed an individual's capacity to reinvent them; models of cultural transmission also show conditions under which complex skills can be lost.[109,110] These findings motivate attention to transmission networks, not a claim that any particular human historical case predicts the fate of an off-Earth community. Archives, training and opportunities for practical participation all matter when expertise must remain usable.

For human-containing systems, social inheritance should preserve the ability to question and change inherited arrangements, not merely obedience to founding instructions. For robotic systems, long-lived operational rules still require identifiable authority, revision procedures and records of consequential decisions. Work on social–ecological systems and polycentric governance offers ways to reason about interacting resource users and institutions.[111,112] Applying those ideas to space is a proposal requiring adaptation, not an empirical result already established for extraterrestrial settlements.

A network of sites could make inheritance useful before any site becomes self-sufficient. One node might specialize in precision measurements, another in bulk processing and another in scientific observation. Shared specifications and redundant services could permit cooperation without identical infrastructure. Knowledge could move more readily than large equipment, while strategic stocks could buffer transport interruptions. But specialization also creates vulnerabilities: the loss of a unique supplier or trusted reference can affect several nodes at once. Network resilience must therefore be tested under correlated failures and contested access, not assumed from the number of locations.

The broader implication is that civilization capabilities might spread by selective re-establishment rather than by exact copying of complete settlements. A new node could inherit only the functions needed for its scientific or infrastructural role and acquire others later. This architecture accommodates robotic precursors, human–machine collaboration and biological–digital combinations without prescribing one inevitable succession. It also leaves room for a deliberate decision not to establish a productive installation where scientific, environmental or social reasons favour restraint.

5.5 Distant exploration as a boundary test

Outer-Solar-System and interstellar scenarios sharpen rather than remove these requirements. As timely external assistance becomes less available, local diagnosis, preservation of knowledge and restoration of key services become more important. Historical self-reproducing-probe proposals remain useful conceptual boundary cases, but their assumed manufacturing and propulsion capabilities are not established by contemporary agent demonstrations.[5] The appropriate next step is to identify the dependencies that survive increasingly severe support restrictions, not assign a date to autonomous galactic expansion.

Models of interstellar settlement explore how transport, site suitability and finite settlement lifetimes can shape spatial patterns.[113] Arguments about technosignatures likewise distinguish the persistence of technological traces from that of their creators.[114] These theoretical perspectives make longevity and propagation scientifically interesting; they are not observations of extraterrestrial civilizations or evidence that human expansion will follow the modelled dynamics. Our framework contributes a complementary question: what physically and institutionally supports the continuation assumed by such models? Exploring the Universe becomes a study of enduring capabilities as well as reachable destinations.

6 Testing and governing the transition: a research agenda

6.1 Measure the connections, not a composite impression of autonomy

The central research problem is compositional. Evidence that agents can plan, robots can manipulate and laboratories can optimize does not establish that their combination will reliably develop new services. Interfaces can introduce incompatible assumptions, error propagation and new demands on supervision. Evaluation should therefore target the transitions between knowledge, realization, qualified use and inheritance. Fig. 5 organizes these tests by scale and evidential gate; Table 3 states the claims and observations that would count against them.

Every experiment should declare the initial system boundary, environment, task distribution, operating interval and allowable external support. Human contributions include advance preparation, demonstrations, interventions, maintenance and post hoc interpretation. Computational support includes model training where relevant, local inference, remote tool calls and emergency access to stronger systems. Resource accounting includes failed trials, discarded material, calibration, tool replacement and reserves consumed. A ground-assisted system may be the best architecture; the requirement is to compare it honestly with alternatives rather than obscure its support.

The resulting evaluation should report a vector of outcomes: qualified service delivered, unsuccessful or unsafe attempts, time to recovery, material and energy use, human effort and dependence on unavailable services. Critical dependencies should be stress-tested individually and in combinations. Removing a remote server while leaving a hidden fallback connection available is not an independence test. Equally, disabling a dependency outside the claimed operating envelope does not refute a narrower claim. Boundaries must be agreed before the result is known.

Statistical discipline is particularly important when physical experiments are expensive. Work on reinforcement-learning evaluation shows how few runs, variable implementations and point estimates can produce misleading rankings.[115,116] Engineering demonstrations should retain unsuccessful campaigns and report uncertainty across tasks, conditions and independent repetitions. A reliability claim cannot be inferred from a single long successful run without considering failure exposure and the population of operating conditions. Detailed operational case studies remain useful, but their evidential role should be stated separately from comparative estimates.

6.2 Three decisive experiments

The first experiment tests *locally developed capability*. An installation receives a service requirement that its qualified initial procedures cannot meet, together with a fixed resource inventory and safety envelope. It may investigate, redesign and fabricate, but acceptance criteria and independent verification are established in advance. The lunar-fixture scenario in Section 3 provides one possible terrestrial analogue. Success requires repeated service after development, not merely a working prototype. Comparators include a fixed automation pipeline, a prequalified library with adaptation, and human-assisted engineering under explicitly costed communication conditions.

The test should distinguish solution retrieval from local development without demanding global novelty. A hidden initial capability, a premanufactured task-specific part or a decisive unreported human correction would weaken the interpretation. So would a result that works only with the exact discovery batch or fails independent inspection. A positive result would show that the system can expand its qualified local repertoire within the declared boundary. It would not yet show that such expansion is economically preferable or sustainable.

The second experiment tests *net scientific–engineering benefit*. Matched installations perform a campaign of useful operations while one is allowed to allocate resources to research and capability development. Baselines can use conventional optimization, maintenance and a fixed procedure library. The comparison tracks cumulative delivered services, downtime, resource depletion and external labour over an interval long enough to expose the costs of development and upkeep. An improvement limited to prototype performance, offset by excessive calibration or maintenance, would not establish net benefit.

This experiment should examine the direction of feedback. Does better characterization produce components that improve later scientific work? Do new instruments provide observations that materially improve subsequent engineering decisions? Or does the system repeatedly rediscover fragile solutions without accumulating reusable knowledge? Withholding the experimental-learning component, the engineering-realization component or the inherited record in separate comparisons can help identify the source of benefit. Where sample sizes are limited, these comparisons should be framed as mechanistic evidence rather than precise universal effect estimates.

The third experiment tests *cross-site inheritance*. A source installation transfers a capability package to an independently configured receiver with different material batches, equipment or environmental conditions. Starting inventories and permitted assistance are recorded. The receiver must establish and sustain an equivalent service under local verification, with adaptation effort and additional imports included. Baselines include documentation alone, transferred software alone and conventional expert-led transfer. Failure despite successful digital copying would locate the missing physical or organizational dependency rather than merely label the receiver insufficiently intelligent.

The strongest version repeats transfer across several sites and after a controlled loss of a supporting function. It asks whether the inherited package contains enough information to diagnose and restore service, not just reproduce a demonstration under ideal conditions. Successful transfer with substantial imported modules remains meaningful if those modules are disclosed. What would weaken the broader inheritance proposition is persistently increasing specialist effort, unresolvable dependence on the source site or loss of independently verified performance after adaptation.

6.3 Trustworthy inference and constrained action

Autonomous development raises a distinctive assurance problem: the system is intended to change what it can do. Safety arguments tied to an initial configuration cannot automatically be extended to its successors. The research agenda should therefore connect uncertainty estimation, constrained control and change management. Bayesian approximations and deep ensembles provide approaches to predictive uncertainty, while conformal prediction supplies coverage guarantees under specified assumptions.[117–119] None supplies unrestricted guarantees under arbitrary environmental change, and confidence estimates should not be treated as authorization to act.

Runtime constraints can provide a separate layer. Control-barrier-function methods formalize restrictions on admissible system behaviour within modelled conditions.[120] Their usefulness depends on the adequacy of the model, sensing and actuation assumptions. A proposed exploration architecture would couple such qualified control layers to agents that recommend actions but cannot silently redefine safety boundaries. If an intended capability requires a new envelope, it should enter an explicit requalification process.

The risks also extend beyond a faulty prediction. Reward misspecification, side effects and unsafe exploration are established concerns in AI safety, while systems-safety research emphasizes hazardous interactions that may occur without an individual component failing.[121,122] For a productive installation, a locally successful optimizer could consume a shared reserve, contaminate a scientific sample or make a component incompatible with another node. Evaluation must therefore include coupled services, adversarial or erroneous tool inputs, and recovery from compromised records. A more accurate model is not by itself a complete systems-safety argument.

6.4 Governance is part of the capability architecture

Technical competence does not determine who may deploy it, alter an environment or establish a new site. The Outer Space Treaty assigns States Parties responsibility for national space activities and requires authorization and continuing supervision of non-governmental activities; Article IX addresses due regard and harmful contamination.[123] COSPAR planetary-protection policy provides a distinct, non-binding guidance framework for protecting scientific investigation and the Earth from relevant contamination risks.[124] Neither should be reduced to a generic label of responsible AI, nor treated as resolving all questions of resource allocation or future social organization.

The engineering implication is that permission must accompany capability. Proposed systems should separate permission to inspect, sample, fabricate, modify infrastructure and establish additional nodes. Local decision-making under delayed communication needs pre-agreed limits, escalation rules and safe fallbacks. A system authorized to maintain one observatory should not infer permission to expand industrial activity because it predicts a scientific benefit. Transfer packages should preserve the scope of authorization and require destination-specific approval where appropriate.

Oversight also needs evidence that remains intelligible across time and institutions. Decision records should connect proposed actions, supporting observations, accepted uncertainty, authorizations and actual outcomes. Independent verification and the ability to suspend a service require resources that must be designed into the installation. For human communities, governance must additionally protect agency, participation and the possibility of revising inherited rules. These are normative commitments of the proposed architecture, not consequences mechanically deduced from engineering performance.

6.5 A programme organized by gates, not dates

The near-term opportunity is a shared terrestrial testbed in which laboratories, robots and maintainable production equipment operate under declared limits on support. Subsequent studies can introduce unfamiliar materials, interrupted communications and environmental qualification before attempting relevant flight demonstrations. Progression should depend on evidence: repeatable new service, positive campaign-level benefit, verified transfer and acceptable oversight. A successful terrestrial result warrants a better-defined next test, not an automatic claim of readiness for the Moon or Mars.

Open capability records, comparable intervention accounting and cross-institutional replications would make these tests cumulative. Space engineers can contribute operational reliability and resource discipline; AI and robotics researchers can contribute adaptable decision and action systems; experimental scientists can contribute valid inference; manufacturing researchers can contribute process qualification; social scientists and affected publics can shape institutional arrangements. The research programme succeeds if it identifies architectures that sustain useful exploration with fewer consequential dependencies—even if the best solution remains a deliberately interdependent human–machine network.

7 Conclusions

The convergence of agentic AI, generalist robotics and autonomous science creates an opportunity to reconsider what is sent into space. Alongside instruments with fixed functions, future missions might carry systems capable of developing additional scientific and engineering services after deployment. Existing research supports important elements of that possibility, but not the assumption that they already form a self-sustaining whole.

Autonomous engineering provides the central connection. It links inquiry to design, physical realization, independent verification and maintained use. Its success should be judged by qualified service and cumulative benefit, not the apparent intelligence of a planner or the novelty of a generated artefact. Energy, feedstocks, production, measurement and computing determine whether those gains persist. Institutions determine which actions are legitimate and which risks are acceptable.

Inheritance extends the question from one installation to many. Digital records can carry knowledge, but a receiving site must re-establish the physical and organizational conditions that make it usable. Mechanical, industrial, biological and social continuity cannot be collapsed into software replication or a single measure of autonomy. Recognizing their differences permits a broader range of architectures, including robotic precursors, human–machine communities and specialized networks that remain intentionally interdependent.

The proposed transition from autonomous science through autonomous engineering to inheritable civilization capabilities is therefore a research hypothesis with observable intermediate claims. Tests of local capability development, net benefit and cross-site re-establishment can reveal both its promise and its limits. The ambition is not to assume that machines will inevitably spread civilization throughout the Universe. It is to understand whether, where and under whose authority the capacity to investigate, build and learn can become durable beyond Earth.

Boxes

Box 1 | Three operational meanings of autonomy

Scientific autonomy is the capacity to select and conduct investigations, interpret evidence and update a scientific account within a declared scope. A platform may be autonomous in choosing experiments while relying on humans for reagent preparation or interpretation of difficult results.

Engineering autonomy is the capacity to turn requirements and evidence into physically realized, independently qualified and maintained services. Planning or fabrication alone does not meet this definition. Nor does behavioural adaptation necessarily constitute physical repair.

Support autonomy is the capacity to retain specified services for a stated interval under stated restrictions on external labour, material, energy, information and computation. It is not a universal property of a robot and should not be expressed without a system boundary.

These dimensions need not rise together. A sophisticated scientific agent connected to Earth-based servers may have substantial scientific autonomy and limited support autonomy. A robust preprogrammed spacecraft may exhibit the opposite pattern. Reports should distinguish human demonstrations from deployment interventions, remote model calls from onboard inference, and inventory consumption from replenishable local production. Planned support is part of the architecture, not automatically a failure. Unreported support makes comparisons uninterpretable.

Box 2 | Why transferring a recipe is not yet inheriting a capability

Suppose site A develops a qualified process for a replacement fixture and sends site B its design, controller and experimental record. Digital transfer is complete when the files arrive intact; capability inheritance is not. Site B must determine whether its feedstock, tools, sensors, processing conditions and acceptance requirements support the same service.

A proposed capability package therefore contains six elements: the service specification and operating envelope; design and executable procedure versions; material and equipment prerequisites; calibration and independent acceptance protocols; failure records and recovery procedures; and the applicable authorization and change-control requirements. Provenance and reproducibility principles support the information layer, but physical re-establishment requires additional tests.[92,93]

The receiver may reproduce, adapt or substitute the original process. Success is established by independent local verification and sustained service over a declared interval, with labour, imports and failed attempts reported. Functional equivalence does not require identical machines. A negative result is informative if it identifies a missing dependency or an invalid transfer assumption. The key distinction is between possessing an instruction and possessing the conditions under which that instruction reliably works.

Table 1 | Representative evidence and the boundary of the demonstrated system

The studies are selected exemplars, not a ranked leaderboard or a complete evidence census. Durations refer only to the reported experiment or operational record, not a common unattended-operation metric. Support statements identify dependencies relevant to interpreting the result; absence of an intervention count is not evidence of zero interventions.

System or study	Demonstrated function	Setting and reported extent	Support boundary	Verification and interpretive limit
AEGIS[60]	Autonomous selection of science targets for instrument observations	Curiosity operations on Mars	Earth-built rover and instruments; ground-defined mission context	Flight science targeting; not new instrument design or local manufacture
Astrobee[62,63]	Free-flying robotic research and assistance; operational fault handling	ISS interior; anomaly report covers more than 750 free-flyer operating hours as of March 2022	Ground updates, crew-supported replacement and return of faulty hardware for repair	Flight operations with documented support; reported hours are not one intervention-free run
Mobile robotic chemist[38]	Closed-loop photocatalysis experimentation	Terrestrial laboratory; 688 experiments over eight days	Existing instruments, prepared research setting and supplied materials	Activity measurements within a bounded campaign; not autonomous laboratory replenishment
A-Lab[54,55]	Robotic synthesis and characterization of inorganic targets	Terrestrial laboratory; reported 17-day campaign	Installed synthesis/characterization equipment and available precursors	Interpret using the corrected article; phase assignment and novelty require qualification
Coscientist[43]	Language-model-guided chemistry with information, code and laboratory tools	Selected terrestrial chemistry tasks	Pre-existing tools and experimental platform	Task and chemistry validation; does not establish independent infrastructure maintenance
AI co-scientist[44]	Hypothesis generation and refinement with specialist-agent collaboration	Selected biomedical applications, including in vitro validation	Scientist-in-the-loop framing, selection and experimental work	Evidence for supported scientific reasoning; not an autonomously operated biological facility
$\pi_{0.5}$[27]	Vision–language–action generalization to new environments	Terrestrial household manipulation evaluations	Co-training data, supplied embodiments and prepared evaluation conditions	Defined task performance; not qualification under planetary conditions

Continued on next page

System or study	Demonstrated function	Setting and reported extent	Support boundary	Verification and interpretive limit
Morphological robot evolution[67]	Robotic construction, testing and iterative design of locomoting assemblies	Controlled terrestrial experiments	Supplied components, assembly infrastructure and evaluation space	Physical design–build feedback; not raw-material closure or field maintenance
ARMADAS[84]	Autonomous assembly of modular structural elements	Metres-scale structures demonstrated in terrestrial gravity	Prefabricated standardized modules and designed assembly interfaces	Ground assembly evidence; lunar infrastructure remains an application vision
MOXIE[79]	Oxygen production from Martian atmospheric carbon dioxide	Mars; cited paper reports seven production runs in 2021	Earth-built instrument integrated with Perseverance and mission support	Planetary resource-processing demonstration; not settlement-scale continuous production or local instrument replacement

Table 2 | Services required to sustain intelligence and production

The service network contains both joint requirements and alternative supply routes. Continuity options are design choices to be compared, not capabilities assumed to exist locally. The relevant metric is qualified service under declared disruptions, not locally produced mass alone.

Service	Principal dependencies	Continuity alternatives	Evidence or design precedent	Required continuity test
Energy and heat rejection	Source, conversion, distribution, storage, radiators or other appropriate thermal interfaces	Redundant sources; protected reserves; load prioritization; replaceable conversion modules	Ground fission-system studies [76,77]	Maintain critical services through a specified generation or cooling shortfall
Feedstock preparation	Characterization, excavation or collection, separation, processing, storage, consumables	Local extraction; recycling; imported specialty inputs; alternative qualified recipes	MOXIE and regolith-processing work [70,79]	Recover acceptable output across batch variation and a disrupted input supply
Manufacturing and assembly	Machines, tooling, fixtures, joining, process control, qualified materials	Fabrication; modular assembly; reversible reconfiguration; stockpiled components	Additive and discrete manufacturing [80,81,83]	Replace a consequential component and demonstrate repeated service, including inspection cost
Metrology and quality assurance	Calibrated instruments, standards, reference materials, independent measurements, traceable records	Redundant principles; protected references; cross-site comparison; periodic external calibration	Proposed extension of qualification and provenance requirements [56]	Detect a planted sensor bias without relying on the same flawed measurement chain
Electronics and sensing	Qualified components, packaging, connectors, power conversion, diagnostic access	Reserves; modular replacement; repair where feasible; selective imports or local production	Radiation mechanisms and assurance literature [86,87]	Restore a required service after a representative electronics failure; disclose replenishment route
Computation and knowledge storage	Power, cooling, processors, memory, storage integrity, software and model compatibility	Efficient local models; verified fallbacks; replicated archives; delayed remote assistance	Compression, distillation and memory-aware computation [89–91]	Operate through loss of remote compute and reconstruct service from an independently checked archive
Logistics and maintenance coordination	Transport, handling, inventories, scheduling, compatible interfaces, authority to exchange	Local buffers; pooled reserves; alternative suppliers; reduced interface diversity	Proposed multi-site synthesis; operational support experience [63]	Sustain priority services through loss of a unique supplier or delayed transport

Table 3 | Questions that make the perspective falsifiable

The tests below are proposed, not performed in this article. Baselines must share declared initial resources or report their differences explicitly. A weakening result constrains the specified proposition and operating envelope; it need not refute every possible architecture.

Question	Proposed test and outcome	Essential comparison	Observation that would weaken the claim
Q1. Which functions can be deployed with less external support?	Run scientific and maintenance tasks with controlled delays, outages and recorded intervention budgets	Fixed autonomy, teleoperation or ground-assisted operation, with support costs stated	Apparent autonomy disappears when hidden preparation, remote computation or rescue actions are counted
Q2. Can a deployed system develop a new local physical capability?	Require a previously unavailable service; verify fabrication, integration and repeated use independently	Prequalified procedure library and expert-assisted adaptation	Only a retrieved design is produced; qualification fails; or an undisclosed task-specific asset is essential
Q3. Does the science–engineering feedback produce net benefit?	Compare cumulative qualified services, resources and downtime with and without capability-development cycles	Matched-resource installations using conventional optimization and maintenance	Development, verification and upkeep consume more resources than the additional service returns
Q4. Can production and computing remain available?	Remove critical support pathways and inject representative faults over a declared interval	Stockpiling, repair, local fabrication and network supply alternatives	High local mass production coexists with immediate failure after loss of one indispensable imported item
Q5. Can capability be inherited across sites?	Transfer a capability package to an independently prepared site with changed conditions; test service locally	Files alone, documentation alone and expert-led transfer	Digital copying succeeds but local service cannot be re-established without increasing specialist dependence
Q6. Can a multi-site system remain governed and resilient?	Exercise authorization changes, disputed resource requests and correlated outages in simulation or controlled analogues	Centralized and distributed arrangements with matched duties and oversight resources	Nodes bypass authority, propagate unverified changes or lose shared critical services despite nominal redundancy

Figures

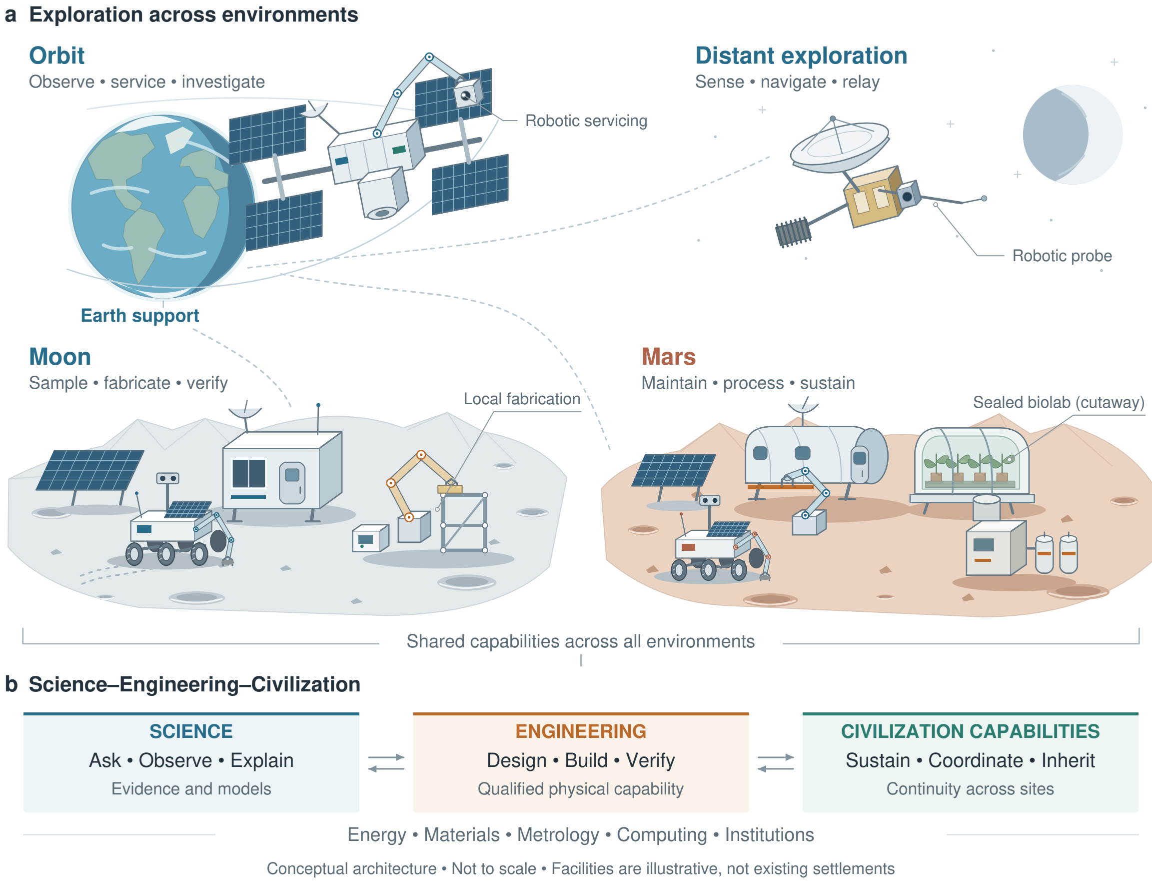

Figure 1: Exploring the Universe through deployable, developing and inheritable capabilities. a, Illustrative roles for agentic AI and robots across orbital, lunar, Martian and distant-exploration settings. Scenes combine instrument servicing, sampling, local fabrication, facility maintenance, material processing and contained biological research. Earth remains an explicit source of support. Dashed links indicate possible cross-environment exchanges, not literal trajectories or presently deployed connections. **b,** A shared framework connects scientific inquiry, verified engineering and the sustained, coordinated inheritance of capabilities. Bidirectional arrows represent feedback, not an inevitable historical progression. Energy, materials, metrology, computing and institutions underpin all three. Activities coexist across settings: no planet corresponds to a single conceptual stage. The original vector illustration is conceptual and not to scale; proposed facilities are not existing settlements, and the scene does not establish self-sufficiency or successful capability inheritance.

Demonstrated functions do not yet form a complete chain

Study-scoped evidence from Table 1 • no summed autonomy or maturity score

	DECISION	REALIZATION			CONTINUITY			
System / evidence	Select inquiry	Physical action	Design / adaptation	Produce / assemble	Independent qualification	Maintain / recover	Cross-site inheritance	Setting / boundary
AEGIS See Table 1	●	●	–	–	–	–	–	Mars flight Autonomous targeting
Astrobee See Table 1	–	●	○	–	–	○	–	ISS flight Ground / crew recovery
Mobile robotic chemist See Table 1	●	●	–	●	–	–	–	Terrestrial laboratory Prepared platform + inputs
A-Lab † See Table 1	●	●	–	●	–	–	–	Terrestrial laboratory Corrected characterization
Coscientist (2023) See Table 1	●	●	–	●	–	–	–	Terrestrial chemistry Installed laboratory tools
AI co-scientist (2026) See Table 1	●	○	–	–	○	–	–	Biomedical studies Scientist-in-the-loop
$\pi_{0.5}$ See Table 1	–	●	–	–	–	–	–	Terrestrial robotics Task-specific evaluation
Morphological evolution See Table 1	–	●	●	●	–	–	–	Terrestrial construction Supplied components
ARMADAS See Table 1	–	●	–	●	–	–	–	Ground demonstration Prefabricated modules
MOXIE See Table 1	–	●	–	●	–	–	–	Mars flight Earth-built processing unit

● Demonstrated within the cited task ○ Externally / human supplied – Not established here

"Produce / assemble" includes synthesis or resource processing; it does not imply hardware manufacture or industrial closure.

Independent qualification is separate from an optimizer's own feedback. Biomedical validation is externally performed; its scope differs from engineering-service acceptance.

† Interpret A-Lab with its published correction. Blank-chain cells are evidence boundaries, not proof that a capability is absent.

Figure 2: Demonstrated functions and unverified connections. The ten rows correspond to Table 1. Filled circles mark task-scoped demonstrations; open circles mark explicitly external or human contributions; short dashes mean that the cited evidence does not establish the function, not that it is impossible or absent. Production includes synthesis and resource processing, not necessarily hardware manufacture. Independent assessment is distinguished from an optimizer's own feedback; the externally performed biomedical validation is not equivalent to engineering-service acceptance. Settings and support boundaries remain visible. A-Lab is interpreted with its published correction. This qualitative first-draft synthesis is not a common benchmark, maturity score or demonstration that the rows compose into one system; cell-level sources accompany the editable figure.

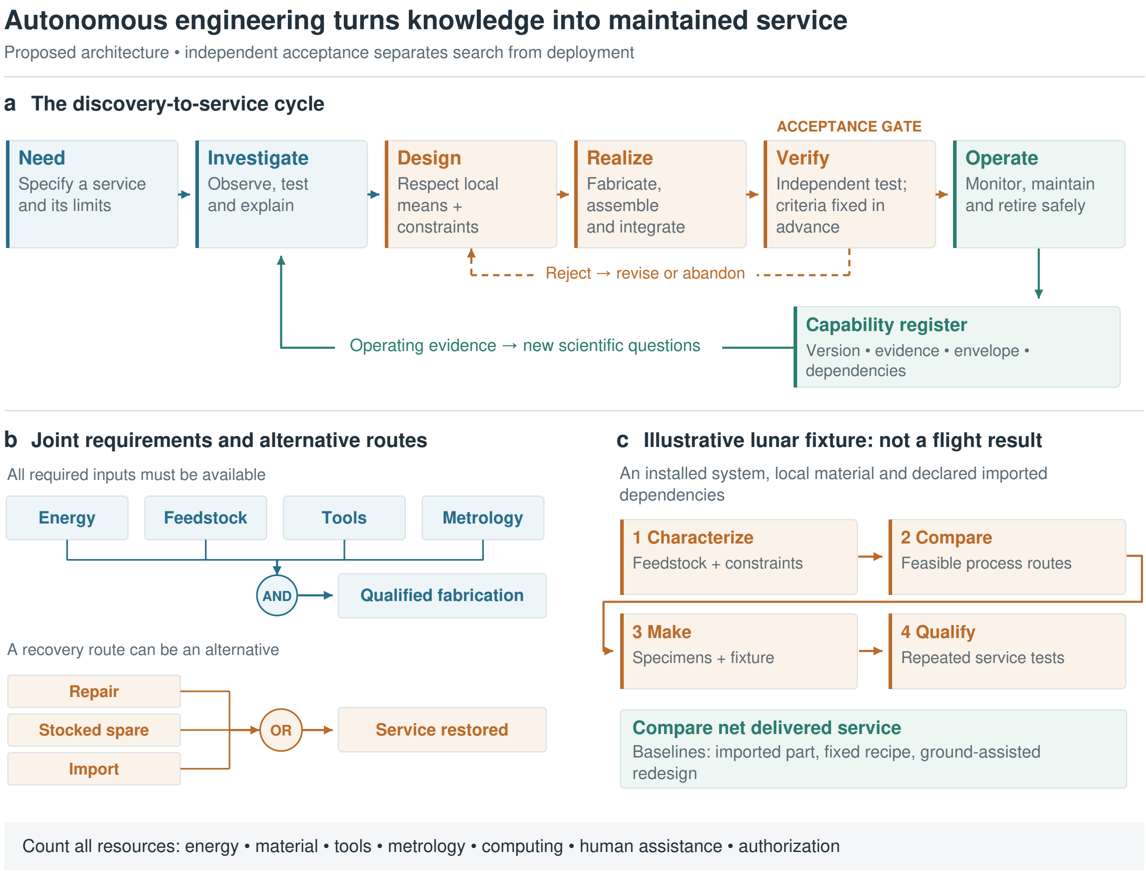

Figure 3: The autonomous engineering bridge. a, A declared service need leads through investigation, locally constrained design, realization and independent verification to maintained operation. Rejected designs return to revision or abandonment; accepted outcomes enter a versioned capability register whose operating evidence informs further inquiry. **b,** AND denotes joint requirements for qualified fabrication; OR denotes alternative recovery routes, conditional on their own prerequisites. The inputs shown are illustrative, not an exhaustive dependency list. **c,** The hypothetical lunar fixture proceeds from characterization through route comparison and fabrication to repeated-service tests. Its benefits must be compared with imported parts, fixed recipes and ground-assisted redesign, including all resource and support costs. The diagram proposes a mechanism and test, not a demonstrated integrated lunar system.

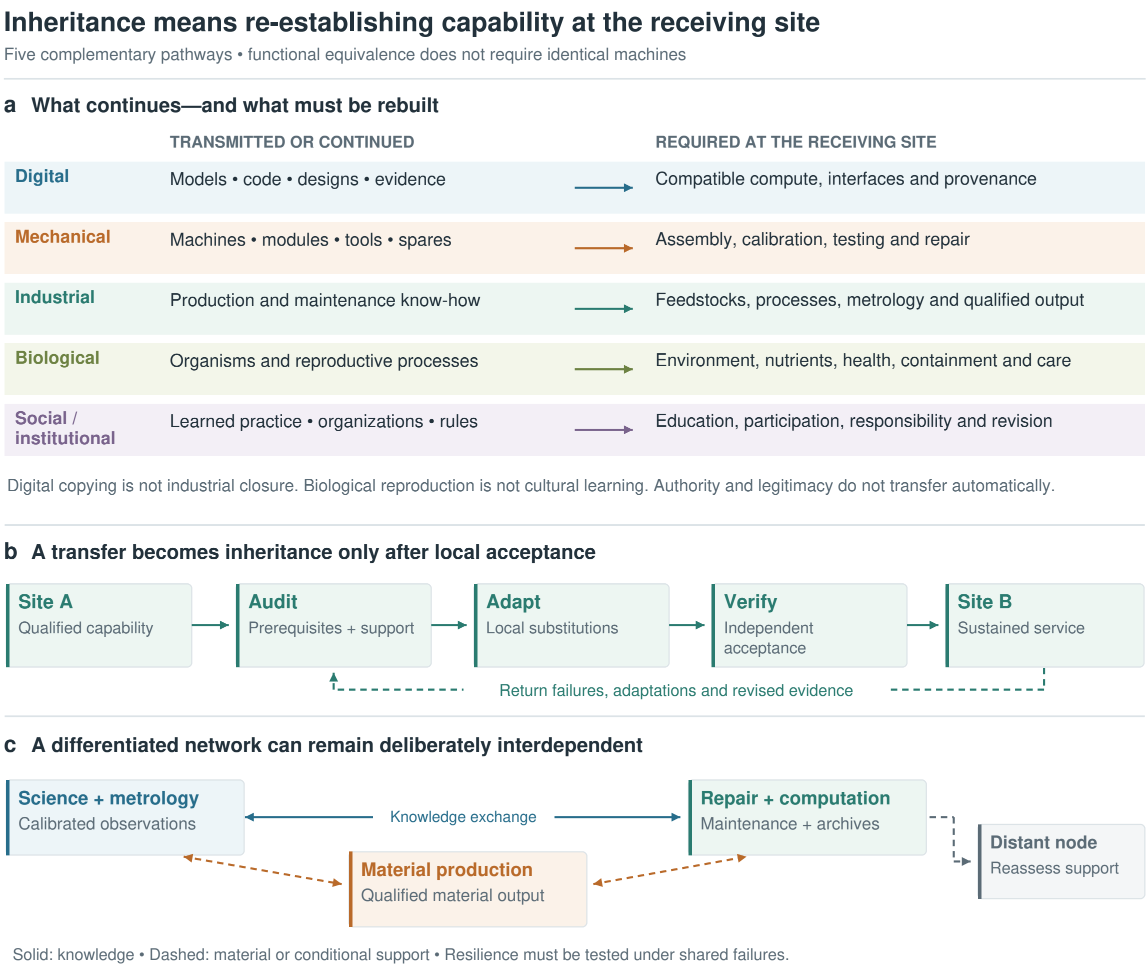

Figure 4: Capability inheritance requires local re-establishment. a, Digital, mechanical, industrial, biological and social/institutional pathways preserve different objects and require complementary support at the receiving site. Authority and legitimacy are not automatically copied. **b,** A transfer from site A becomes an inherited service at site B only after prerequisite audit, adaptation and independent local acceptance; failures and revised evidence feed back into the record. **c,** A proposed differentiated network exchanges knowledge through solid links and material or conditional support through dashed links. More distant nodes require their support assumptions to be reassessed. These exchanges are architectural possibilities, not proven resilience: shared failures and interruptions must be tested. Not every node is assumed to contain people or biological production.

Advance by evidence gates, not by a calendar forecast

Proposed tests across research contexts • all cells are research designs, not completed results

	Component studies	Integrated terrestrial facility	Environmental qualification	Flight-relevant context	Multi-site system
Q2 **Local capability**	New tool; independent test	Closed discovery-to-service loop	Changed feedstock or operating conditions	New service within an approved envelope	Recipient establishes a compatible service
Q3 **Net benefit**	Useful output versus full cost	Campaign-level gain; matched resources	Benefit persists after environmental change	Mission service versus external support burden	Network benefit versus coordination cost
Q5 **Inheritance**	Recalibration on new hardware	Transfer across independently prepared workshops	Requalify under different local conditions	Re-establish after declared support loss	Cross-site transfer; independent local acceptance

a Shared evidential requirements

- **Declare support** — Human work + remote compute
- **Verify independently** — Acceptance outside the optimizer
- **Account over time** — Failures, upkeep + service delivered
- **Stress dependencies** — Outages, faults + recovery

b Governance remains a distinct condition

- Permission to act
- Contamination control
- Change audit + oversight
- Human agency + participation

Tests can run in parallel. A positive result narrows the next question; it does not imply readiness for more distant deployment.

Figure 5: An evidence-gated research programme. Columns specify research contexts, not dates or a compulsory sequence. Rows organize proposed tests of locally developed capability (Q2), net scientific–engineering benefit (Q3) and cross-site inheritance (Q5), as defined in Table 3. **a,** Shared evidential requirements include declared support, independent acceptance, campaign-level accounting and dependency stress tests. **b,** Governance is a separate condition spanning permission, contamination control, change oversight and human agency where relevant. Tests may proceed in parallel; success motivates a more precisely bounded next experiment rather than automatic readiness for a more distant destination. No cell reports a completed experiment.

References

[1] Gao, Y. & Chien, S. Review on space robotics: Toward top-level science through space exploration. *Science Robotics* **2**, eaan5074 (2017). https://doi.org/10.1126/scirobotics.aan5074.

[2] Yao, S. *et al.* ReAct: Synergizing Reasoning and Acting in Language Models. arXiv manuscript (2022). https://arxiv.org/abs/2210.03629.

[3] Open X-Embodiment Collaboration *et al.* Open X-Embodiment: Robotic Learning Datasets and RT-X Models. arXiv manuscript (2023). https://arxiv.org/abs/2310.08864.

[4] Tom, G. *et al.* Self-Driving Laboratories for Chemistry and Materials Science. *Chemical Reviews* **124**, 9633–9732 (2024). https://doi.org/10.1021/acs.chemrev.4c00055.

[5] Freitas, R. A., Jr. A Self-Reproducing Interstellar Probe. *Journal of the British Interplanetary Society* **33**, 251–264 (1980). https://www.rfreitas.com/Astro/ReproJBISJuly1980.htm. Linked author-hosted version is an earlier draft and may differ from the published article.

[6] Metzger, P. T., Muscatello, A., Mueller, R. P. & Mantovani, J. Affordable, Rapid Bootstrapping of the Space Industry and Solar System Civilization. *Journal of Aerospace Engineering* **26**, 18–29 (2013). https://doi.org/10.1061/(asce)as.1943-5525.0000236. **Connects contemporary autonomy to the earlier proposal of progressively bootstrapping off-Earth industry.**

[7] Schick, T. *et al.* Toolformer: Language Models Can Teach Themselves to Use Tools. arXiv manuscript (2023). https://arxiv.org/abs/2302.04761.

[8] Shinn, N. *et al.* Reflexion: Language Agents with Verbal Reinforcement Learning. arXiv manuscript (2023). https://arxiv.org/abs/2303.11366.

[9] Wang, G. *et al.* Voyager: An Open-Ended Embodied Agent with Large Language Models. arXiv manuscript (2023). https://arxiv.org/abs/2305.16291.

[10] DeepSeek-AI *et al.* DeepSeek-R1: Incentivizing Reasoning Capability in LLMs via Reinforcement Learning. arXiv manuscript (2025). https://arxiv.org/abs/2501.12948.

[11] Novikov, A. *et al.* AlphaEvolve: A coding agent for scientific and algorithmic discovery. arXiv manuscript (2025). https://arxiv.org/abs/2506.13131.

[12] Jimenez, C. E. *et al.* SWE-bench: Can Language Models Resolve Real-World GitHub Issues? arXiv manuscript (2023). https://arxiv.org/abs/2310.06770.

[13] Yang, J. *et al.* SWE-agent: Agent-Computer Interfaces Enable Automated Software Engineering. arXiv manuscript (2024). https://arxiv.org/abs/2405.15793.

[14] Liu, X. *et al.* AgentBench: Evaluating LLMs as Agents. arXiv manuscript (2023). https://arxiv.org/abs/2308.03688.

[15] Zhou, S. *et al.* WebArena: A Realistic Web Environment for Building Autonomous Agents. arXiv manuscript (2023). https://arxiv.org/abs/2307.13854.

[16] Xie, T. *et al.* OSWorld: Benchmarking Multimodal Agents for Open-Ended Tasks in Real Computer Environments. arXiv manuscript (2024). https://arxiv.org/abs/2404.07972.

[17] Kwa, T. *et al.* Measuring AI Ability to Complete Long Software Tasks. arXiv manuscript (2025). https://arxiv.org/abs/2503.14499.

[18] Ahn, M. *et al.* Do As I Can, Not As I Say: Grounding Language in Robotic Affordances. arXiv manuscript (2022). https://arxiv.org/abs/2204.01691.

[19] Liang, J. *et al.* Code as Policies: Language Model Programs for Embodied Control. arXiv manuscript (2022). https://arxiv.org/abs/2209.07753.

[20] Driess, D. *et al.* PaLM-E: An Embodied Multimodal Language Model. arXiv manuscript (2023). https://arxiv.org/abs/2303.03378.

[21] Brohan, A. *et al.* RT-1: Robotics Transformer for Real-World Control at Scale. arXiv manuscript (2022). https://arxiv.org/abs/2212.06817.

[22] Brohan, A. *et al.* RT-2: Vision-Language-Action Models Transfer Web Knowledge to Robotic Control. arXiv manuscript (2023). https://arxiv.org/abs/2307.15818.

[23] Octo Model Team *et al.* Octo: An Open-Source Generalist Robot Policy. arXiv manuscript (2024). https://arxiv.org/abs/2405.12213.

[24] Kim, M. J. *et al.* OpenVLA: An Open-Source Vision-Language-Action Model. arXiv manuscript (2024). https://arxiv.org/abs/2406.09246.

[25] Chi, C. *et al.* Diffusion Policy: Visuomotor Policy Learning via Action Diffusion. arXiv manuscript (2023). https://arxiv.org/abs/2303.04137.

[26] Black, K. *et al.* π_0: A Vision-Language-Action Flow Model for General Robot Control. arXiv manuscript (2024). https://arxiv.org/abs/2410.24164.

[27] Intelligence, P. *et al.* $\pi_{0.5}$: a Vision-Language-Action Model with Open-World Generalization. arXiv manuscript (2025). https://arxiv.org/abs/2504.16054.

[28] Gemini Robotics Team *et al.* Gemini Robotics 1.5: Pushing the Frontier of Generalist Robots with Advanced Embodied Reasoning, Thinking, and Motion Transfer. arXiv manuscript (2025). https://arxiv.org/abs/2510.03342.

[29] Zhao, T. Z., Kumar, V., Levine, S. & Finn, C. Learning Fine-Grained Bimanual Manipulation with Low-Cost Hardware. arXiv manuscript (2023). https://arxiv.org/abs/2304.13705.

[30] Fu, Z., Zhao, T. Z. & Finn, C. Mobile ALOHA: Learning Bimanual Mobile Manipulation with Low-Cost Whole-Body Teleoperation. arXiv manuscript (2024). https://arxiv.org/abs/2401.02117.

[31] Fu, Z., Zhao, Q., Wu, Q., Wetzstein, G. & Finn, C. HumanPlus: Humanoid Shadowing and Imitation from Humans. arXiv manuscript (2024). https://arxiv.org/abs/2406.10454.

[32] Lan, Z. *et al.* AutoBio: A Simulation and Benchmark for Robotic Automation in Digital Biology Laboratory. arXiv manuscript (2025). https://arxiv.org/abs/2505.14030.

[33] King, R. D. *et al.* Functional genomic hypothesis generation and experimentation by a robot scientist. *Nature* **427**, 247–252 (2004). https://doi.org/10.1038/nature02236.

[34] King, R. D. *et al.* The Automation of Science. *Science* **324**, 85–89 (2009). https://doi.org/10.1126/science.1165620.

[35] Granda, J. M., Donina, L., Dragone, V., Long, D.-L. & Cronin, L. Controlling an organic synthesis robot with machine learning to search for new reactivity. *Nature* **559**, 377–381 (2018). https://doi.org/10.1038/s41586-018-0307-8.

[36] Coley, C. W. *et al.* A robotic platform for flow synthesis of organic compounds informed by AI planning. *Science* **365**, eaax1566 (2019). https://doi.org/10.1126/science.aax1566.

[37] Steiner, S. *et al.* Organic synthesis in a modular robotic system driven by a chemical programming language. *Science* **363**, eaav2211 (2019). https://doi.org/10.1126/science.aav2211.

[38] Burger, B. *et al.* A mobile robotic chemist. *Nature* **583**, 237–241 (2020). https://doi.org/10.1038/s41586-020-2442-2. **Demonstrates an extended physical experiment-selection loop in a pre-existing laboratory.**

[39] MacLeod, B. P. *et al.* Self-driving laboratory for accelerated discovery of thin-film materials. *Science Advances* **6**, eaaz8867 (2020). https://doi.org/10.1126/sciadv.aaz8867.

[40] Nikolaev, P. *et al.* Autonomy in materials research: a case study in carbon nanotube growth. *npj Computational Materials* **2**, 16031 (2016). https://doi.org/10.1038/npjcompumats.2016.31.

[41] Attia, P. M. *et al.* Closed-loop optimization of fast-charging protocols for batteries with machine learning. *Nature* **578**, 397–402 (2020). https://doi.org/10.1038/s41586-020-1994-5.

[42] Shields, B. J. *et al.* Bayesian reaction optimization as a tool for chemical synthesis. *Nature* **590**, 89–96 (2021). https://doi.org/10.1038/s41586-021-03213-y.

[43] Boiko, D. A., MacKnight, R., Kline, B. & Gomes, G. Autonomous chemical research with large language models. *Nature* **624**, 570–578 (2023). https://doi.org/10.1038/s41586-023-06792-0.

[44] Gottweis, J. *et al.* Accelerating scientific discovery with Co-Scientist. *Nature* **655**, 487–496 (2026). https://doi.org/10.1038/s41586-026-10644-y. **Illustrates scientist-in-the-loop agent collaboration and selected experimental validation.**

[45] Jumper, J. *et al.* Highly accurate protein structure prediction with AlphaFold. *Nature* **596**, 583–589 (2021). https://doi.org/10.1038/s41586-021-03819-2.

[46] Merchant, A. *et al.* Scaling deep learning for materials discovery. *Nature* **624**, 80–85 (2023). https://doi.org/10.1038/s41586-023-06735-9.

[47] Fawzi, A. *et al.* Discovering faster matrix multiplication algorithms with reinforcement learning. *Nature* **610**, 47–53 (2022). https://doi.org/10.1038/s41586-022-05172-4.

[48] Trinh, T. H., Wu, Y., Le, Q. V., He, H. & Luong, T. Solving olympiad geometry without human demonstrations. *Nature* **625**, 476–482 (2024). https://doi.org/10.1038/s41586-023-06747-5.

[49] Romera-Paredes, B. *et al.* Mathematical discoveries from program search with large language models. *Nature* **625**, 468–475 (2024). https://doi.org/10.1038/s41586-023-06924-6.

[50] Mankowitz, D. J. *et al.* Faster sorting algorithms discovered using deep reinforcement learning. *Nature* **618**, 257–263 (2023). https://doi.org/10.1038/s41586-023-06004-9.

[51] Schmidt, M. & Lipson, H. Distilling Free-Form Natural Laws from Experimental Data. *Science* **324**, 81–85 (2009). https://doi.org/10.1126/science.1165893.

[52] Brunton, S. L., Proctor, J. L. & Kutz, J. N. Discovering governing equations from data by sparse identification of nonlinear dynamical systems. *Proceedings of the National Academy of Sciences* **113**, 3932–3937 (2016). https://doi.org/10.1073/pnas.1517384113.

[53] Udrescu, S.-M. & Tegmark, M. AI Feynman: A physics-inspired method for symbolic regression. *Science Advances* **6**, eaay2631 (2020). https://doi.org/10.1126/sciadv.aay2631.

[54] Szymanski, N. J. *et al.* An autonomous laboratory for the accelerated synthesis of inorganic materials. *Nature* **624**, 86–91 (2023). https://doi.org/10.1038/s41586-023-06734-w.

[55] Szymanski, N. J. *et al.* Author Correction: An autonomous laboratory for the accelerated synthesis of inorganic materials. *Nature* **650**, E1–E1 (2026). https://doi.org/10.1038/s41586-025-09992-y. **Shows why phase identification and novelty claims require an independently defensible evidential chain.**

[56] Canty, R. B. & Abolhasani, M. The past, present and future of self-driving laboratories. *Nature Reviews Chemistry* **10**, 523–537 (2026). https://doi.org/10.1038/s41570-026-00847-2.

[57] Wang, H. *et al.* Scientific discovery in the age of artificial intelligence. *Nature* **620**, 47–60 (2023). https://doi.org/10.1038/s41586-023-06221-2.

[58] Muscettola, N., Nayak, P., Pell, B. & Williams, B. C. Remote Agent: to boldly go where no AI system has gone before. *Artificial Intelligence* **103**, 5–47 (1998). https://doi.org/10.1016/s0004-3702(98)00068-x.

[59] Chien, S. *et al.* Using Autonomy Flight Software to Improve Science Return on Earth Observing One. *Journal of Aerospace Computing, Information, and Communication* **2**, 196–216 (2005). https://doi.org/10.2514/1.12923.

[60] Francis, R. *et al.* AEGIS autonomous targeting for ChemCam on Mars Science Laboratory: Deployment and results of initial science team use. *Science Robotics* **2**, eaan4582 (2017). https://doi.org/10.1126/scirobotics.aan4582.

[61] Nesnas, I. A. *et al.* Axel and DuAxel rovers for the sustainable exploration of extreme terrains. *Journal of Field Robotics* **29**, 663–685 (2012). https://doi.org/10.1002/rob.21407.

[62] Smith, T. *et al.* Astrobee: A New Platform for Free-Flying Robotics on the International Space Station. In *Proceedings of i-SAIRAS 2016*, 20160007769 (2016). https://ntrs.nasa.gov/citations/20160007769.

[63] Carlino, R., Barlow, J., Benavides, J., Katterhagen, A. & Cortez, J. Recovering from On-orbit Anomalies on the Astrobee Free Flyers and its Systems. NASA Technical Reports Server (2022). https://ntrs.nasa.gov/citations/20220005920. Presentation, ISS Research and Development Conference, 2022; NTRS 20220005920.

[64] Krenn, M., Malik, M., Fickler, R., Lapkiewicz, R. & Zeilinger, A. Automated Search for new Quantum Experiments. *Physical Review Letters* **116**, 090405 (2016). https://doi.org/10.1103/physrevlett.116.090405.

[65] Klimesch, J. *et al.* Designing physics experiments with artificial intelligence. *Nature* **657**, 47–58 (2026). https://doi.org/10.1038/s41586-026-10898-6.

[66] Lipson, H. & Pollack, J. B. Automatic design and manufacture of robotic lifeforms. *Nature* **406**, 974–978 (2000). https://doi.org/10.1038/35023115.

[67] Brodbeck, L., Hauser, S. & Iida, F. Morphological Evolution of Physical Robots through Model-Free Phenotype Development. *PLOS ONE* **10**, e0128444 (2015). https://doi.org/10.1371/journal.pone.0128444. **Connects computational design to robotic construction and physical evaluation within a bounded system.**

[68] Stella, F., Della Santina, C. & Hughes, J. How can LLMs transform the robotic design process? *Nature Machine Intelligence* **5**, 561–564 (2023). https://doi.org/10.1038/s42256-023-00669-7.

[69] Schwandt, C., Hamilton, J. A., Fray, D. J. & Crawford, I. A. The production of oxygen and metal from lunar regolith. *Planetary and Space Science* **74**, 49–56 (2012). https://doi.org/10.1016/j.pss.2012.06.011.

[70] Lomax, B. A. *et al.* Proving the viability of an electrochemical process for the simultaneous extraction of oxygen and production of metal alloys from lunar regolith. *Planetary and Space Science* **180**, 104748 (2020). https://doi.org/10.1016/j.pss.2019.104748.

[71] Meurisse, A., Makaya, A., Willsch, C. & Sperl, M. Solar 3D printing of lunar regolith. *Acta Astronautica* **152**, 800–810 (2018). https://doi.org/10.1016/j.actaastro.2018.06.063.

[72] Cesaretti, G., Dini, E., De Kestelier, X., Colla, V. & Pambaguian, L. Building components for an outpost on the Lunar soil by means of a novel 3D printing technology. *Acta Astronautica* **93**, 430–450 (2014). https://doi.org/10.1016/j.actaastro.2013.07.034.

[73] Bongard, J., Zykov, V. & Lipson, H. Resilient Machines Through Continuous Self-Modeling. *Science* **314**, 1118–1121 (2006). https://doi.org/10.1126/science.1133687.

[74] Cully, A., Clune, J., Tarapore, D. & Mouret, J.-B. Robots that can adapt like animals. *Nature* **521**, 503–507 (2015). https://doi.org/10.1038/nature14422.

[75] Degrave, J. *et al.* Magnetic control of tokamak plasmas through deep reinforcement learning. *Nature* **602**, 414–419 (2022). https://doi.org/10.1038/s41586-021-04301-9.

[76] McClure, P. R., Poston, D. I., Gibson, M. A., Mason, L. S. & Robinson, R. C. Kilopower Project: The KRUSTY Fission Power Experiment and Potential Missions. *Nuclear Technology* **206** (2020). https://doi.org/10.1080/00295450.2020.1722554.

[77] Poston, D. I., Gibson, M. A., Sanchez, R. G. & McClure, P. R. Results of the KRUSTY Nuclear System Test. *Nuclear Technology* **206** (2020). https://doi.org/10.1080/00295450.2020.1730673.

[78] Hecht, M. *et al.* Mars Oxygen ISRU Experiment (MOXIE). *Space Science Reviews* **217**, 9 (2021). https://doi.org/10.1007/s11214-020-00782-8.

[79] Hoffman, J. A. *et al.* Mars Oxygen ISRU Experiment (MOXIE)—Preparing for human Mars exploration. *Science Advances* **8**, eabp8636 (2022). https://doi.org/10.1126/sciadv.abp8636. **Establishes resource processing on Mars while leaving infrastructure scaling and replacement as separate questions.**

[80] DebRoy, T. *et al.* Additive manufacturing of metallic components – Process, structure and properties. *Progress in Materials Science* **92**, 112–224 (2018). https://doi.org/10.1016/j.pmatsci.2017.10.001.

[81] Prater, T. *et al.* 3D Printing in Zero G Technology Demonstration Mission: complete experimental results and summary of related material modeling efforts. *The International Journal of Advanced Manufacturing Technology* **101**, 391–417 (2019). https://doi.org/10.1007/s00170-018-2827-7.

[82] Cheung, K. C. & Gershenfeld, N. Reversibly Assembled Cellular Composite Materials. *Science* **341**, 1219–1221 (2013). https://doi.org/10.1126/science.1240889.

[83] Jenett, B., Abdel-Rahman, A., Cheung, K. & Gershenfeld, N. Material–Robot System for Assembly of Discrete Cellular Structures. *IEEE Robotics and Automation Letters* **4**, 4019–4026 (2019). https://doi.org/10.1109/lra.2019.2930486.

[84] Gregg, C. E. & Cheung, K. C. Automated Reconfigurable Mission Adaptive Digital Assembly Systems (ARMADAS): Robotically Assembled Sustainable Lunar Infrastructure. NASA Technical Reports Server (2023). https://ntrs.nasa.gov/citations/20230001316. Presentation; NTRS 20230001316. Ground demonstration and lunar infrastructure concept.

[85] Baumann, R. Radiation-induced soft errors in advanced semiconductor technologies. *IEEE Transactions on Device and Materials Reliability* **5**, 305–316 (2005). https://doi.org/10.1109/tdmr.2005.853449.

[86] Dodd, P. & Massengill, L. Basic mechanisms and modeling of single-event upset in digital microelectronics. *IEEE Transactions on Nuclear Science* **50**, 583–602 (2003). https://doi.org/10.1109/tns.2003.813129.

[87] Schwank, J. R., Shaneyfelt, M. R. & Dodd, P. E. Radiation Hardness Assurance Testing of Microelectronic Devices and Integrated Circuits: Radiation Environments, Physical Mechanisms, and Foundations for Hardness Assurance. *IEEE Transactions on Nuclear Science* **60**, 2074–2100 (2013). https://doi.org/10.1109/tns.2013.2254722.

[88] Horowitz, M. 1.1 Computing's energy problem (and what we can do about it). In *2014 IEEE International Solid-State Circuits Conference Digest of Technical Papers (ISSCC)*, 10–14 (2014). https://doi.org/10.1109/isscc.2014.6757323.

[89] Han, S., Mao, H. & Dally, W. J. Deep Compression: Compressing Deep Neural Networks with Pruning, Trained Quantization and Huffman Coding. arXiv manuscript (2015). https://arxiv.org/abs/1510.00149.

[90] Hinton, G., Vinyals, O. & Dean, J. Distilling the Knowledge in a Neural Network. arXiv manuscript (2015). https://arxiv.org/abs/1503.02531.

[91] Dao, T., Fu, D. Y., Ermon, S., Rudra, A. & Ré, C. FlashAttention: Fast and Memory-Efficient Exact Attention with IO-Awareness. arXiv manuscript (2022). https://arxiv.org/abs/2205.14135.

[92] Wilkinson, M. D. *et al.* The FAIR Guiding Principles for scientific data management and stewardship. *Scientific Data* **3**, 160018 (2016). https://doi.org/10.1038/sdata.2016.18.

[93] Peng, R. D. Reproducible Research in Computational Science. *Science* **334**, 1226–1227 (2011). https://doi.org/10.1126/science.1213847.

[94] Zykov, V., Mytilinaios, E., Adams, B. & Lipson, H. Self-reproducing machines. *Nature* **435**, 163–164 (2005). https://doi.org/10.1038/435163a.

[95] Jones, R. *et al.* RepRap – the replicating rapid prototyper. *Robotica* **29**, 177–191 (2011). https://doi.org/10.1017/s026357471000069x. **Provides a concrete precedent for partial machine reconstruction whose external dependencies remain essential.**

[96] Montague, M. *et al.* The Role of Synthetic Biology for In Situ Resource Utilization (ISRU). *Astrobiology* **12**, 1135–1142 (2012). https://doi.org/10.1089/ast.2012.0829.

[97] Menezes, A. A., Cumbers, J., Hogan, J. A. & Arkin, A. P. Towards synthetic biological approaches to resource utilization on space missions. *Journal of The Royal Society Interface* **12**, 20140715 (2015). https://doi.org/10.1098/rsif.2014.0715.

[98] Nangle, S. N. *et al.* The case for biotech on Mars. *Nature Biotechnology* **38**, 401–407 (2020). https://doi.org/10.1038/s41587-020-0485-4.

[99] Verseux, C. *et al.* Sustainable life support on Mars – the potential roles of cyanobacteria. *International Journal of Astrobiology* **15**, 65–92 (2016). https://doi.org/10.1017/s147355041500021x.

[100] Verseux, C. *et al.* A Low-Pressure, N2/CO2 Atmosphere Is Suitable for Cyanobacterium-Based Life-Support Systems on Mars. *Frontiers in Microbiology* **12**, 611798 (2021). https://doi.org/10.3389/fmicb.2021.611798.

[101] Clauwaert, P. *et al.* Nitrogen cycling in Bioregenerative Life Support Systems: Challenges for waste refinery and food production processes. *Progress in Aerospace Sciences* **91**, 87–98 (2017). https://doi.org/10.1016/j.paerosci.2017.04.002.

[102] Wakayama, S. *et al.* Healthy offspring from freeze-dried mouse spermatozoa held on the International Space Station for 9 months. *Proceedings of the National Academy of Sciences* **114**, 5988–5993 (2017). https://doi.org/10.1073/pnas.1701425114.

[103] Wakayama, S. *et al.* Evaluating the long-term effect of space radiation on the reproductive normality of mammalian sperm preserved on the International Space Station. *Science Advances* **7**, eabg5554 (2021). https://doi.org/10.1126/sciadv.abg5554.

[104] Wakayama, S. *et al.* Effect of microgravity on mammalian embryo development evaluated at the International Space Station. *iScience* **26**, 108177 (2023). https://doi.org/10.1016/j.isci.2023.108177.

[105] Garrett-Bakelman, F. E. *et al.* The NASA Twins Study: A multidimensional analysis of a year-long human spaceflight. *Science* **364**, eaau8650 (2019). https://doi.org/10.1126/science.aau8650.

[106] Afshinnekoo, E. *et al.* Fundamental Biological Features of Spaceflight: Advancing the Field to Enable Deep-Space Exploration. *Cell* **183**, 1162–1184 (2020). https://doi.org/10.1016/j.cell.2020.10.050.

[107] Kriegman, S., Blackiston, D., Levin, M. & Bongard, J. A scalable pipeline for designing reconfigurable organisms. *Proceedings of the National Academy of Sciences* **117**, 1853–1859 (2020). https://doi.org/10.1073/pnas.1910837117.

[108] Kriegman, S., Blackiston, D., Levin, M. & Bongard, J. Kinematic self-replication in reconfigurable organisms. *Proceedings of the National Academy of Sciences* **118**, e2112672118 (2021). https://doi.org/10.1073/pnas.2112672118.

[109] Boyd, R., Richerson, P. J. & Henrich, J. The cultural niche: Why social learning is essential for human adaptation. *Proceedings of the National Academy of Sciences* **108**, 10918–10925 (2011). https://doi.org/10.1073/pnas.1100290108.

[110] Henrich, J. Demography and Cultural Evolution: How Adaptive Cultural Processes Can Produce Maladaptive Losses—The Tasmanian Case. *American Antiquity* **69**, 197–214 (2004). https://doi.org/10.2307/4128416.

[111] Ostrom, E. A General Framework for Analyzing Sustainability of Social-Ecological Systems. *Science* **325**, 419–422 (2009). https://doi.org/10.1126/science.1172133.

[112] Ostrom, E. Polycentric systems for coping with collective action and global environmental change. *Global Environmental Change* **20**, 550–557 (2010). https://doi.org/10.1016/j.gloenvcha.2010.07.004.

[113] Carroll-Nellenback, J., Frank, A., Wright, J. & Scharf, C. The Fermi Paradox and the Aurora Effect: Exo-civilization Settlement, Expansion, and Steady States. *The Astronomical Journal* **158**, 117 (2019). https://doi.org/10.3847/1538-3881/ab31a3.

[114] Wright, J. T. *et al.* The Case for Technosignatures: Why They May Be Abundant, Long-lived, Highly Detectable, and Unambiguous. *The Astrophysical Journal Letters* **927**, L30 (2022). https://doi.org/10.3847/2041-8213/ac5824.

[115] Henderson, P. *et al.* Deep Reinforcement Learning that Matters. arXiv manuscript (2017). https://arxiv.org/abs/1709.06560.

[116] Agarwal, R., Schwarzer, M., Castro, P. S., Courville, A. & Bellemare, M. G. Deep Reinforcement Learning at the Edge of the Statistical Precipice. arXiv manuscript (2021). https://arxiv.org/abs/2108.13264.

[117] Gal, Y. & Ghahramani, Z. Dropout as a Bayesian Approximation: Representing Model Uncertainty in Deep Learning. arXiv manuscript (2015). https://arxiv.org/abs/1506.02142.

[118] Lakshminarayanan, B., Pritzel, A. & Blundell, C. Simple and Scalable Predictive Uncertainty Estimation using Deep Ensembles. arXiv manuscript (2016). https://arxiv.org/abs/1612.01474.

[119] Angelopoulos, A. N. & Bates, S. A Gentle Introduction to Conformal Prediction and Distribution-Free Uncertainty Quantification. arXiv manuscript (2021). https://arxiv.org/abs/2107.07511.

[120] Ames, A. D., Xu, X., Grizzle, J. W. & Tabuada, P. Control Barrier Function Based Quadratic Programs for Safety Critical Systems. *IEEE Transactions on Automatic Control* **62**, 3861–3876 (2017). https://doi.org/10.1109/tac.2016.2638961.

[121] Amodei, D. *et al.* Concrete Problems in AI Safety. arXiv manuscript (2016). https://arxiv.org/abs/1606.06565.

[122] Leveson, N. A new accident model for engineering safer systems. *Safety Science* **42**, 237–270 (2004). https://doi.org/10.1016/s0925-7535(03)00047-x. **Frames safety in terms of system interactions and constraints rather than isolated component reliability.**

[123] United Nations. Treaty on Principles Governing the Activities of States in the Exploration and Use of Outer Space, including the Moon and Other Celestial Bodies. United Nations Office for Outer Space Affairs (1967). https://www.unoosa.org/oosa/en/ourwork/spacelaw/treaties/outerspacetreaty.html?stream=science. Articles VI and IX.

[124] Committee on Space Research. COSPAR Policy on Planetary Protection. COSPAR (2026). https://cosparhq.cnes.fr/assets/uploads/2026/01/SRT_224_Editorial_PPP-Nov-2025.pdf. Approved 7 November 2025; published January 2026.