Grey Matters
ExploreGreyMatters.ai

CASE STUDY • PUBLICATION 01

The Things We Can Now Attempt

What a year of curiosity, practical projects and human–AI collaboration taught us

FROM THE UNIVERSE TO THE BACK GARDEN

Human-led. AI-enabled.

Grey Matters • October 2026 • Published edition v1.5

It started with a conversation

I wasn't setting out to build a website, develop a research programme or write a book about the universe. Nor was I looking for a new way to plan a holiday, design a garden or choose a camper van. I was curious about things that mattered to me, large and small.

One starting point was science, especially the difficult bits involving cosmology and quantum physics. But the attraction of AI wasn't limited to science. It was the possibility of talking through almost any question, from something baffling in a book to an ordinary problem at home.

AI offered an intriguing possibility. Instead of reading something I didn't understand and moving on, I could ask about it. Then question the explanation. Then ask for a simpler explanation, a different perspective or an example.

Occasionally I could even explain it to George, our West Highland Terrier. His grasp of quantum mechanics remains difficult to assess, but it maybe better than mine!

Something was changing. AI was no longer simply a quicker way to find information. It was becoming something I could think with.

What else might I attempt if I didn't have to become an expert in everything before I could begin?

That question eventually became Grey Matters.

From the universe to the back garden

Curiosity has a habit of getting out of hand. A question about physics led to quantum mechanics, cosmology and an increasingly speculative attempt to understand how things might fit together. Meanwhile, entirely different conversations were helping with travel, a garden, creative ideas and practical decisions. There was no master plan, and no requirement to know how to do everything before starting.

The science conversations became a substantial manuscript, then a physical book, conceptual diagrams and visualisations. Later, after challenge and revision, they became a more carefully bounded conceptual framework called Cosmogenic Selection. We had moved from reading difficult ideas to attempting to explain and critically examine some of our own.

This didn't mean we had discovered how the universe works. Being able to attempt something is not the same as establishing that it is correct. But it suggested that AI's value might be about more than speed: it could change the altitude of what we aimed to achieve, the depth of our exploration and the range of subjects we felt able to approach.

At the other end of the scale, the same conversational approach was useful for ordinary things. Instead of asking AI for a finished answer, we could explain what mattered, explore options, question assumptions and adjust the result until it became useful.

The subject could change completely; the starting point remained simple. Bring a real interest or problem, have a conversation and keep the judgement human.

Planning a three-week trip to Hawaii became an exercise in bringing together places, interests, costs, routes, accommodation and the inevitable complications of travel. The interesting part wasn't simply producing an itinerary. It was the conversation around it: what we really wanted to experience, what we'd rather avoid, where the plan was becoming too ambitious and what mattered enough to justify changing it.

Similarly, a garden that began as part of a new-build property became a real design and construction project. Ideas became visual possibilities, alternatives could be compared, and the final result could be discussed before anyone started moving earth.

Our camper-van search introduced practical constraints: vehicle length, passengers, storage, electric bicycles, weight limits, cost and the difference between something attractive in a showroom and something that would actually work for us. AI could help assemble and compare choices. It couldn't decide what mattered most to us.

Other projects were more imaginative. An illustrated creature called Snolphin began with a child's imagination and developed into pictures and a little world of its own. The creative spark belonged to the child; AI helped explore where it might lead.

And then there was Bitcoin. Rather than asking AI to predict its price, we used it to examine the proposition, competing views and a question that remains useful long after a particular prediction expires:

What would make us change our minds?

Across these experiences a pattern emerged. The useful starting point wasn't the technology. It was something we cared about. The human brought purpose, experience, preferences, uncertainties and judgement. AI brought breadth, speed, alternatives, explanation and the ability to help produce something tangible.

The results weren't always correct, and not every conversation produced something worth keeping. But we were increasingly able to attempt things that had previously seemed difficult, expensive or beyond our reach.

When an interesting answer isn't enough

Some projects revealed a less comfortable truth: AI can produce a beautifully coherent explanation of something that isn't necessarily true. In cosmology, this is particularly serious. An attractive narrative can be satisfying while remaining speculative. A computer simulation can behave exactly as programmed without demonstrating that nature behaves the same way.

We learned to separate four questions. Does the thing work? Does it reproduce what we expect? Does it represent what we think it represents? And does the evidence justify the conclusion?

A financial model can calculate every number correctly while misunderstanding the economic consequence of a transaction. A simulation can accurately demonstrate a proposed mechanism without establishing that the mechanism exists in the physical universe. An explanation can be internally consistent while overlooking a decisive alternative.

Our Cosmogenic Selection work therefore evolved into a research programme designed to challenge the idea, not protect it. Different AI perspectives examined claims through different disciplinary lenses. Findings were reconciled without hiding disagreement. A separate auditing role challenged conclusions that went beyond the evidence, and a human decided what could change in the approved research record.

Crucially, no change was a legitimate outcome. One reviewed proposition became more precisely defined without acquiring stronger evidential support. That was useful progress, even if it didn't make the theory more impressive.

The research eventually reached external publication. Publication made the ideas available for examination; it did not establish scientific validity. Nor are AI systems playing specialist roles substitutes for qualified human researchers, or their agreement scientific consensus.

Collaboration becomes more valuable when AI is encouraged to challenge our thinking, not merely help us defend it.

The lesson also applies to investment decisions, reports and other work where a confident explanation might conceal a weak assumption.

The danger of enjoying yourself too much

This collaboration can be absorbing. On holiday, I could look out at the sea with a book and my phone nearby, exploring an idea, refining an image or creating something that previously required specialist skills and tools. It was liberating. It was also occasionally intrusive.

One answer led to another. One image suggested an improvement. One project opened another. Sometimes the technical machinery needed to produce an apparently simple result demanded more attention than the result justified.

The boundary between an enjoyable intellectual pursuit and an unexpected operational commitment was surprisingly easy to cross. That became especially apparent when we built a public website.

The accidental release operator

By October 2026 Grey Matters had accumulated books, research papers, visualisations, practical projects, decision frameworks, images and lessons worth sharing. A website seemed a sensible place to bring them together.

It would also test the collaboration: could someone without a software-development background use AI to design, build and operate a credible public website? The answer turned out to be yes, but with considerable qualifications.

AI was remarkably capable at designing pages, developing architecture, writing code, explaining unfamiliar systems and identifying problems. Reasoning about a task and completing it across different services, however, were very different things.

Accounts needed authorising. Permissions didn't always work. Browser sessions crashed. Long conversations lost their place. One successful operation appeared to have failed because an

automated check misunderstood the response. At other times, genuinely faulty changes looked plausible until separately tested.

I spent extraordinary amounts of time navigating dashboards, copying technical identifiers between windows, recovering interrupted work and trying to establish what had actually happened.

The irony was hard to miss. We set out to show how AI could enable a non-specialist to accomplish more. Somehow I became the project's accidental release operator.

Some involvement was essential: deciding scope, controlling privileged access and determining what became public. But copying strings between systems and repeatedly pressing buttons wasn't a valuable use of human judgement. Our division of responsibility needed to improve.

When more becomes less

AI also makes ambitious ideas remarkably easy to propose. A simple way for visitors to contact Grey Matters grew into structured submissions, a database, authentication, administrative controls and associated engineering. Each component had a rationale; together they created complexity our small, unproven audience hadn't requested.

More features required more testing, environments and permissions. The project became increasingly elaborate while the website itself remained unpublished.

An external publication deadline helped force a decision. What did visitors actually need on day one? A way to explore the work and contact us. An email address would do nicely.

We preserved the more ambitious engineering as a future option, reduced immediate scope and launched the website on 22 October 2026.

The ability to build something is not, on its own, a reason to build it.

Sometimes the best progress is to stop improving the thing that hasn't yet been delivered.

A method we didn't know we were developing

Looking back, the same discipline had emerged across very different projects. In learning it helped distinguish understanding from accepting an explanation. In planning it turned preferences into workable choices. In decisions it exposed assumptions and what might change our minds. In research it separated hypotheses from evidence. In engineering it prevented experimental changes from silently replacing what already worked.

We summarised it in five instructions:

Explore freely. Persist deliberately. Build deliberately. Verify adversarially. Freeze decisively.

Explore protects the pleasure of discovery. Persist means keeping important decisions outside a chat that can end or lose context. Build means making a defined change or creating a usable output while protecting the approved starting point. Verify means searching for errors and

checking the real result, not merely trusting a reassuring account. Freeze means a person accepts an identifiable result with its limitations.

The discipline needn't look identical everywhere. A book discussion might need nothing beyond conversation. An important plan may benefit from a short record of preferences and assumptions. Publishing a research claim, altering a financial model or allowing AI to act on a live website requires much stronger checks.

We call these levels Conversational, Structured and Governed. Start lightly; add structure when consequences justify it. And learn the reverse lesson too: once work is sufficiently checked, stop adding controls simply because more controls are possible.

The point wasn't to eliminate the human

After launch, further improvements provided a test of a better division of work. Instead of operating every deployment system ourselves, we prepared bounded procedures that an AI execution agent could run within defined limits. The agent could prepare changes, run checks and report evidence; it couldn't legitimately expand its permissions or substitute a different release when something unexpected occurred.

On one release a permission rule blocked the agent. It stopped and reported the problem, rather than finding a way round. That was precisely what we wanted. The human adjusted the permission boundary and the controlled process resumed.

The site received further public improvements, including prompts visitors could use in their own AI conversations. We haven't systematically counted the difference in manual effort, and we haven't established that the same approach works equally well for other people and platforms.

What we observed in our circumstances was that the human could concentrate increasingly on purpose, judgement, acceptance and decisions that mattered, while AI and automation performed more repeatable work between them.

The goal wasn't human absence. It was human involvement where human judgement adds value.

What might somebody else take from this?

Grey Matters isn't evidence that anyone can become an expert at anything simply by talking to AI. Nor does it prove that every task becomes faster, every answer can be trusted or every ambitious idea deserves pursuing.

It suggests something more practical: you don't necessarily need to master every discipline and production tool before you begin exploring a worthwhile goal. You do need to know what you want, recognise gaps in your knowledge, challenge the answers and retain responsibility for consequential decisions.

Begin with a difficult book, family project, design idea or important choice. Ask AI to help you understand rather than simply answer. Give it your priorities. Invite criticism. If the

conversation grows into a project, keep a short decision record. If it affects other people, money, systems or public claims, check more carefully.

You needn't adopt our whole method. Most people do not need an Evidence Observatory to organise a holiday. Neither does George.

Where we've landed

Grey Matters began with curiosity and led to a book, scientific exploration, imaginative creations, practical decisions, analytical frameworks and a public website. Some things worked remarkably well. Some were difficult. Some ideas were abandoned or narrowed when the evidence didn't justify them.

The most interesting contribution wasn't simply answering questions faster. It was expanding what we felt able to investigate, create and undertake, while also requiring us to improve how we question, check, decide and sometimes stop.

Our experience is not a universal recipe. It comes from one continuing collaboration; its wider usefulness remains to be tested. But it offers an invitation.

Start with something that matters to you. Have a conversation. Challenge what comes back. See where it leads.

You might surprise yourself. And if you spend a holiday discovering why a website deployment thinks it has failed when it actually succeeded, you may have taken things a little further than intended.

Try it yourself: the companion starter playbook

If this story has sparked an idea, the companion Grey Matters guide, Working with AI on What Matters to You, turns the lessons into seven ways to begin, short prompts to adapt, everyday examples and simple checks. You can use it without reading this case study — and without knowing anything about how websites are built.

Grey Matters — Human-led. AI-enabled.

The subject is not the transferable asset. The collaboration is.