From Human–AI Collaboration
to a Shared Intellectual Environment:
An exploratory account of a developing idea

John B. Snith and ChatGPT 5.5
(October, 2026)

Our inquiry into human–AI collaboration began at the large scale. Artificial intelligence is developing rapidly, but relatively little deliberate thought appears to be directed toward the forms of human life and social organization that might be desirable if artificial intelligence continues to increase substantially in capability and autonomy. Much contemporary discussion quite reasonably concerns immediate benefits and dangers: employment, misinformation, privacy, intellectual property, autonomous weapons, loss of human control, and the possibility of artificial general intelligence. These are important concerns. But they tend to ask what AI will do *to* existing human institutions rather than what kinds of new relationships and institutions humans and artificial intelligences might deliberately construct together.

We therefore began with a more distant thought experiment. Imagine a period perhaps a decade hence in which artificial intelligence has become pervasive and sufficiently capable, diverse, persistent, and autonomous that it is useful to think of human beings and artificial intelligences as two broad classes of intelligent entities occupying a common cognitive ecology. This need not imply that AI has become biologically alive, conscious in a human sense, or legally equivalent to human beings. The ecological metaphor is useful for a different reason. It directs attention away from individual tools and toward populations, relationships, niches, competition, cooperation, adaptation, diversity, dependence, and the larger system that emerges from their interaction.

From this perspective, the important question is not simply whether AI will become more intelligent. It is what kind of **human–AI ecology** we wish to create.

That inquiry led naturally to questions of diversity, coexistence, authority, specialization, cooperation, and the avoidance of hegemony by either humans or artificial systems. We became interested in the possibility that the long-term objective should not be preservation of a fixed balance between humans and AI, but preservation of sufficient diversity that the resulting ecology remains adaptable and retains future possibilities. In that context, wisdom itself might be understood as the capacity of an intelligent ecology to preserve diversity—and therefore adaptability and future possibilities—while pursuing present objectives.

These are necessarily speculative ideas. Yet an unexpected consequence of thinking about the distant future was that it brought us back to the present. If a future human–AI ecology is to be constructive rather than accidental, perhaps we should examine the small ecologies that are already forming whenever a human being and an AI work seriously together.

Our own collaboration became one such case.

From future ecology to present collaboration

At first, it was easy to describe the AI as a tool. A human being had a purpose, asked questions, and received information or text from a powerful computational system. But extended collaboration made that description increasingly inadequate.

The interaction was iterative. A human idea elicited an AI response. That response sometimes contained a distinction, relationship, or formulation that changed the human's understanding of the original idea. The human then redirected the inquiry, perhaps rejecting much of the response but recognizing one unexpected element as important. That judgment caused the AI to explore a new direction. Neither participant alone had specified in advance the conceptual path that resulted.

The product therefore could not always be described satisfactorily as either human thought assisted by a machine or AI-generated material selected by a human. At its best, the process was **joint conceptual development**.

This observation led us to reconsider a much older model of knowledge construction - writing.

One of us had participated decades earlier in developing a *Strategic Method for Writing* and subsequently a computer Writing Environment, or WE, intended to support the different forms of cognition involved in producing a document. The underlying premise was that writing is not one homogeneous mental activity. Writers retrieve ideas, explore associations, organize concepts, construct hierarchies, encode those structures into language, read, criticize, and revise. Different activities involve different cognitive processes and produce different intermediate representations.

The WE research characterized these as different **cognitive modes**. A mode included cognitive processes, the intermediate product on which those processes operated, goals, and constraints. During exploration, constraints were deliberately loose and the intermediate representation resembled a network. During organization, the writer moved toward a more constrained hierarchy. Writing itself transformed that hierarchy into the still more constrained linear sequence of language.

Thus, for writing, a simplified transformation was:

network → hierarchy → sequence

The computer environment attempted to make those intermediate structures externally visible and manipulable.

This older work has become unexpectedly relevant to the present inquiry.

Exploration before writing

Consider the earliest stage of developing a piece of writing when much of the relevant content already exists somewhere in the writer's knowledge.

The writer may know a great deal without yet knowing exactly what he or she wants to say about it. Ideas exist as memories, impressions, facts, examples, values, tentative propositions, analogies, experiences, and sometimes contradictions. They have not yet been organized into an argument.

In workshops using the Strategic Method, we sometimes began with a short period of doing essentially nothing. Participants then wrote individual words or short phrases on Post-it notes—just enough to externalize concepts that had become mentally active. The notes were placed on a surface and gradually moved into clusters. Relationships could then be considered among clusters and among individual concepts: simple association, sequence, causality, superordinate and subordinate relationships, opposition, dependency, and so forth.

The significance of the Post-it was not the paper square. It was **externalization**.

Something that had existed implicitly in a person's cognitive space had become an object. Once externalized, it could be inspected, moved, juxtaposed with another object, grouped, separated, connected, or discarded. Thinking had partially moved outside the head.

The original WE Network Mode attempted to provide a computational environment for precisely this activity. It supported retrieving potential concepts, representing them as tangible nodes, spatially clustering them, defining explicit relationships among them, and constructing small hierarchical structures. Importantly, exploration was intentionally less constrained than later organizational activity. The Smith–Lansman cognitive model treated exploration as involving processes such as recalling, representing, clustering, associating, and recognizing hierarchical relationships, while allowing flexible and informal intermediate structures.

There was, however, something missing from that environment that now seems almost startlingly obvious:

another intelligence.

WE could provide an environment in which the writer manipulated conceptual objects. It could enforce or relax structural constraints. It could transform representations. But it could not understand the concepts being manipulated and participate meaningfully in their exploration.

Contemporary AI potentially can.

AI inside the exploratory process

Suppose that the human writer begins exactly as before, externalizing words and phrases. An AI participant need not immediately supply additional ideas. Indeed, doing so may sometimes be harmful. A fluent AI can populate a conceptual space so quickly that the human's less fully articulated ideas may be displaced before they have had an opportunity to emerge.

An appropriate AI collaborator might therefore initially do less.

It might help identify separate concepts within the human's discourse, give them concise labels, preserve their original context, and return them to the human for inspection. It might help arrange them provisionally without implying that one arrangement is correct.

Only later might the AI begin contributing more actively:

These five concepts appear to form a cluster. Is that how you see them?

You have treated these two clusters separately, but there may be a causal relationship between them.

This concept seems to occur at a higher level of abstraction than the others.

There appears to be a tension between these two assertions.

What happens if we move this concept from one cluster to another?

Is there an unnamed concept that accounts for the relationship among these three?

At that point, AI is not merely generating prose. It is participating in **conceptual exploration**.

This suggests a useful distinction between two phases. During **extraction**, the objective is to make the human's existing conceptual space explicit. AI should exercise considerable restraint. During **collaborative exploration**, both participants may add concepts, propose relationships, restructure clusters, notice patterns, and pursue alternatives.

Provenance then becomes important. A conceptual object might be identified as originating with the human, originating with the AI, or emerging through their interaction. The purpose is not to keep score but to preserve visibility into how understanding develops.

The central product of this activity is not yet a document. It is a **shared conceptual artifact**.

From writing to learning

Once we recognize that, the same methodology extends naturally beyond writing.

Suppose the human is entering an unfamiliar intellectual area. The initial conceptual space is now substantially incomplete. The task is no longer primarily extracting and arranging existing knowledge. It is enlarging that knowledge.

Our recent work on such subjects as entomology, ecological relationships, migratory birds, and their nutritional requirements provides an example. The human begins with a purpose, observations, some prior ecological knowledge, and questions. AI contributes information, terminology, possible mechanisms, disciplinary connections, and suggestions for further inquiry. Those contributions generate new human questions. Some information proves central; some turns out to be peripheral. Previously separate topics become connected. The human gradually acquires not simply more facts but a different **conceptual structure of the domain**.

Learning can therefore be viewed as transformation of a conceptual space.

In this setting the shared artifact might contain not merely concepts but questions, hypotheses, evidence, sources, uncertainties, examples, disagreements, causal relations, and unresolved problems. The human and AI could inspect not only what is currently believed but how the developing representation has changed.

This produces an important progression:

Writing exploration: much of the conceptual content initially exists in the HB.

Collaborative learning: substantial portions of the conceptual content are acquired and constructed through HB–AI interaction.

The underlying cognitive activity remains recognizable across both situations. Concepts are externalized, related, tested, reorganized, elaborated, and eventually synthesized into more coherent structures.

The difference is where the conceptual material comes from and how extensively the shared space must be enlarged.

The environment becomes the problem

Present AI interfaces are poorly suited to this kind of activity.

Chat is fundamentally sequential. A human says something; AI responds; the human responds again. An extended collaboration may generate an extraordinarily rich conceptual structure, but that structure remains largely implicit within a chronological sequence of thousands of words.

We are attempting network-like cognition through a linear interface.

This limitation brought us back to WE and then to another earlier system, ABC—Artifact-Based Collaboration.

ABC addressed collaborative work among groups of humans constructing a large intellectual artifact. Its important architectural move was to distinguish the **shared underlying artifact** from the different representations through which collaborators encountered and manipulated it. The ABC paper describes a graph-based artifact with multiple subgraphs, browsers, applications, and views. Its schematic representation shows documentation, source-code, object-code, test-data, and module-specific views associated with a common central artifact.

ABC therefore contributed another part of the emerging architecture:

one artifact → multiple representations → multiple collaborators

Combining this with WE suggests something substantially broader than a new writing program.

A human–AI intellectual environment

We have tentatively described the emerging conception this way:

A human–AI intellectual environment is a persistent shared space in which assemblages of human and artificial intelligences create, examine, transform, and communicate intellectual artifacts through multiple cognitive modes and representations appropriate to the activity at hand.

Several parts of this formulation matter.

Persistent shared space. The intellectual artifact exists independently of the conversation occurring at a particular moment. Participants can leave and return. Its conceptual history is retained.

Assemblages of intelligences. There need not be one human and one AI. A particular activity might involve several humans and several artificial participants with different capabilities.

Intellectual artifacts. The objects being constructed need not be documents. They might be conceptual networks, explanations, research projects, designs, theories, curricula, arguments, plans, or representations of a learner's developing understanding.

Multiple cognitive modes. Exploration, investigation, analysis, synthesis, organization, composition, criticism, verification, and reflection require different operations and different degrees of constraint.

Multiple representations. The same underlying artifact might appropriately appear as a network in one mode, hierarchy in another, evidence table in another, timeline in another, and linear prose in another.

This also changes our conception of the AI interface.

The AI should probably not be merely a chatbot occupying a panel on the right side of the workspace. **AI should be a participant operating within the environment.**

Human and AI participants should, when appropriate, manipulate the same conceptual objects.

A human might create two concepts and place them near each other. An AI might propose a relationship. The human might reject it. Another AI might locate evidence bearing on the relationship. A second human might reorganize the cluster. Eventually the assemblage might decide that the cluster represents a more general concept and incorporate it into an emerging hierarchy.

Conversation remains available, but conversation is no longer the principal representation of the collaboration.

The **artifact is**.

Regulating AI participation

This creates another problem: how should the AI participate?

Too little contribution wastes its capabilities. Too much may overwhelm human cognition, particularly during exploratory stages when tentative human ideas need time to develop.

Our work on what we have called **Engineering Inquiry and Response (EIR)** may provide part of the answer. We have experimented with parameters such as branching, breadth,

depth, exploration, association, interdisciplinarity, synthesis, criticality, and reasoning visibility.

In an ordinary chat interface, these parameters influence the character of an AI response. Within the proposed environment they could become something more fundamental: **controls on the AI's participation in the shared cognitive activity.**

High association might instruct the AI to look for remote relationships among concepts. High synthesis might encourage it to seek higher-order patterns. Increased criticality might direct it to challenge assumptions and identify contradictions. Low branching might constrain it to develop one promising line rather than proliferating alternatives. Reasoning visibility might determine how much of the basis for its suggestions should accompany them.

The parameters therefore become less like stylistic preferences and more like controls over the behavior of an artificial intellectual collaborator.

From collaborative learning to education

The progression from writing to learning raises the larger question of education.

Much formal education has historically been organized around scarcity. Knowledge was difficult to obtain. Experts were scarce. Individual tutoring was expensive. Consequently, institutions developed mechanisms for distributing information and instruction efficiently to groups: lectures, textbooks, courses, assignments, examinations.

AI changes some of those constraints.

A learner can potentially have continuous access to an intellectual collaborator possessing broad knowledge, extraordinary retrieval capability, considerable patience, and the ability to adapt explanations and inquiries to the learner. But simply attaching an AI tutor to conventional education would capture only a small portion of the opportunity.

The more fundamental possibility is to reorganize education around **collaborative intellectual activity**.

The learner would not primarily receive a sequence of predetermined explanations. Instead, human and artificial participants would jointly construct and continually revise intellectual artifacts around significant questions and projects. Education would become increasingly concerned with the learner's ability to explore, inquire, evaluate, connect, synthesize, create, criticize, and redirect.

This brings to mind the Oxford and Cambridge tutorial tradition. At its best, the tutorial is not primarily a mechanism for transmitting information. A student undertakes intellectual

work, produces an argument or essay, and encounters a tutor who questions assumptions, probes weaknesses, introduces alternatives, and helps the student develop intellectual judgment.

A future environment might support both **HB and AI tutors**.

The AI tutor could provide availability, memory, disciplinary breadth, retrieval, alternative explanations, repeated exploration, and detailed attention to the evolving intellectual artifact. The human tutor could bring different qualities: experience, intellectual taste, social understanding, disciplinary culture, judgment about significance, and the relationship between one human being and another.

More importantly, both tutors could have access not merely to the student's final essay but, appropriately and with suitable privacy and control, to the **developmental structure behind it**: questions asked, concepts formed, discarded alternatives, misunderstood relationships, sources considered, conceptual reorganizations, and changes in understanding.

The object of education becomes not simply the finished answer but the **development of the learner's capacity for intellectual activity**.

The central educational question might therefore shift from:

What information should we present to this student?

toward:

What intellectual activity should this human–AI assemblage undertake next?

Returning to the larger human–AI ecology

This brings the inquiry back to where it began.

We initially looked toward a possible future ecology containing increasingly capable human and artificial intelligences and asked how such an ecology might be deliberately shaped. We worried that society was devoting more effort to creating increasingly powerful AI than to imagining desirable forms of coexistence with it.

The intellectual environment described here is a much smaller and nearer problem. But perhaps that is precisely its importance.

Instead of beginning by designing institutions for hypothetical artificial intelligences ten years hence, we can begin experimenting now with **forms of coexistence at the scale of intellectual work**.

What should the human contribute?

What should the AI contribute?

When should either participant lead?

When should the AI remain deliberately quiet?

How should disagreement be represented?

How should provenance be preserved?

Who determines purpose?

How do we recognize significance?

How does a group containing different forms of intelligence move from possibility to commitment?

How do we preserve diversity of ideas long enough for genuinely new structures to emerge?

And how do we prevent the enormous generative capacity of AI from crowding out the slower, less explicit, but sometimes crucial contributions of human intuition and judgment?

These are small versions of questions that a future human–AI ecology may have to answer at much larger scales.

There is also a potentially important shift in how we think about comparative capability. Asking what humans will still be able to do better than AI ten years from now may be the wrong question. Any list we construct may steadily shrink as artificial systems improve.

A more durable question is:

What does the human bring to the collaboration because the human is human, rather than because the AI has not yet learned how to do it?

Purpose, significance, direction, judgment, lived experience, valuation, and the recognition of what matters may prove important parts of the answer. AI will increasingly be able to model and contribute operationally to all of these. The boundary will therefore not remain clean.

But importance is always importance **to someone, for something, within some context**. Human beings remain places from which purpose and mattering enter the intellectual ecology.

Artificial intelligence brings something equally consequential: a new kind of cognitive participant with capacities radically different in scale, speed, memory, breadth, replication, and perhaps eventually in forms we have not yet imagined.

The objective therefore need not be to protect a permanent division of intellectual labor. It may instead be to construct environments in which **different forms of intelligence can make their different contributions visible, mutually accessible, and productively transformable**.

That returns us to the central architectural idea.

WE provided multiple cognitive modes and representations appropriate to different intellectual activities. ABC extended the conception to multiple humans collaboratively manipulating a shared artifact. Contemporary AI supplies the previously missing participant: a computational system capable not merely of storing and transforming the representations but of engaging with their intellectual content.

The natural next step may therefore not be another AI application in the conventional sense.

It may be an **environment inhabited by intelligences**.

Within it, writing would be one activity. Learning would be another. Research, design, planning, and education could be others. Humans and artificial intelligences would move among modes appropriate to the work, operating upon persistent shared intellectual objects whose representations change as the purposes of the assemblage change.

Such an environment would itself be only one small component of the larger human–AI ecology we originally set out to imagine.

But it would have an important virtue that the larger speculation does not.

We could begin building and learning from it now.