

Agentic Media: Reimagining the Future of Communication

YUN WANG and YAN LU, Microsoft Research Asia, China

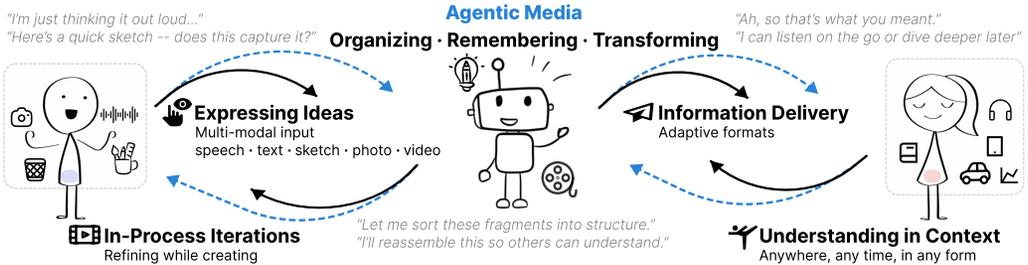


Fig. 1. From raw multimodal expression, semantic structuring, to contextual comprehension, agentic media reframes communication as an iterative and adaptive process

Traditional media systems model communication as a linear transfer of static content from author to reader, mediated by largely passive tools. This model enforces rigid separations between creation and consumption, fragments communicative processes, and constrains collaboration as meaning evolves over time. We introduce Agentic Media, a communication paradigm in which media participate in the construction and negotiation of meaning by embedding communicative intent, retaining interactional context, and supporting adaptive engagement. Within this paradigm, communication is reframed as an ongoing process of expression, exploration, interpretation, and reflection, rather than the delivery of finalized artifacts. We articulate the conceptual foundations for reasoning about how communication can be structured when media are treated as active participants in communicative processes. Building on this framing, we outline research opportunities for HCI and AI that shift attention from artifact manipulation toward communication-centric computing. This work aims to broaden how computing supports sustained communication, collaboration, and meaning-making among people in AI-enabled environments.

Additional Key Words and Phrases: Agentic Media, Human-AI Collaboration, Interactive Media Systems, Knowledge Work, Digital Media, Agentic Memory, Agentic Authoring, Adaptive Interface, Human-AI Co-Creation, Human-Agent Interaction

1 Introduction

Communication has always been the foundation of human civilization. It is through communication that knowledge spreads, collaboration scales, and societies sustain collective memory across generations. Historically, this capacity has relied on stable, physical media, such as printed texts that enabled durable dissemination of ideas, or broadcast technologies that carried information rapidly across the time and distance.

These media fundamentally shaped how knowledge was shared and preserved, while also scaffolding new forms of organization, coordination, and cultural development. As digital technologies emerged, early computer software and user interfaces inherited this tradition. To make novel tools intuitive, they emulated familiar physical metaphors: word processors mirrored paper pages, presentation software replicated slides, and spreadsheets reproduced the tabular grids of accounting ledgers [9, 12, 31, 64]. In all these cases, digital technologies extended the long-standing tradition of treating communication as the transfer of finalized content stored in static containers.

Over time, this format-centric organization has introduced structural constraints into contemporary knowledge work. Formats do not merely influence how information is expressed; they shape how work is divided, coordinated, and understood. In collaborative settings, different formats privilege different operations and forms of reasoning, encouraging contributors to specialize around particular tools and representational logics [3, 35, 53]. As projects grow in scale and complexity, this coupling between roles and formats becomes increasingly pronounced. Work fragments across specialized tools and workflows, isolating contributors within local contexts and making it harder to sustain shared understanding across the broader project [25, 45].

The rapid rise of generative artificial intelligence appears, at first glance, to ease the fragmentation imposed by format-bound media. Large language models and multimodal systems can readily translate across representational boundaries, reassembling materials that were previously scattered across documents, slides, spreadsheets, and other forms, suggesting the possibility of more fluid and integrated communication. However, this promise is rarely realized in practice. Typical uses of AI today often give rise to absurd cycles of communication. Imagine a sender typing a simple, single-sentence prompt, from which the AI produces extensive, detailed content. At the receiving end, another AI model aggressively condenses this output back into a similarly short summary. At first glance, both sides might feel they have gained substantial information. In reality, this superficial cycle of over-generation followed by over-compression dilutes or distorts the original communicative intent [34, 63]. Far from improving clarity or deepening understanding, such loops waste computational resources, squander human attention, and ultimately undermine rather than enhance meaningful dialogue—exposing the limits of treating communication as the manipulation of finalized artifacts [10, 59].

At its core, the breakdown described above reflects a misalignment between how media artifacts are structurally treated and how they are used in practice. While contemporary digital infrastructures continue to assume documents, slides, and messages as finalized endpoints, people increasingly use generative AI to treat these artifacts as transient intermediates to be repeatedly transformed and re-expressed rather than read as final outcomes.

We stand at a technological inflection point where existing media structures increasingly strain to sustain how communication unfolds in practice. As generative AI makes it increasingly easy to transform, re-express, and circulate content, media infrastructures, along with the software and professional practices built around them, remain organized around static, finalized artifacts and increasingly struggle to support continuity, intent, and coordination. These conditions call for a rethinking of media themselves.

Against this backdrop, we use the term **Agentic Media** to articulate a different way of thinking about digital media and communication (Figure 2). In Agentic Media, communication is treated as an operable and adaptive process that unfolds over time, instead of a sequence of finalized artifacts exchanged across formats. Within such media, the medium itself participates in communication in situated and responsive ways, engaging with users to support the ongoing formation, interpretation, and revision of meaning. The conceptual framing of Agentic Media is intentionally open, allowing it to evolve alongside changes in technology, practice, and communicative form. Taken together, this paper clarifies why format-bound media structures are increasingly misaligned with contemporary communication practices and develops Agentic Media as a conceptual framing for rethinking the role of media under these conditions. Rather than offering a closed theory or a technical proposal, the paper aims to reframe how communication is organized, coordinated, and sustained over time, providing a foundation for future exploration across interaction design, collaborative systems, and media infrastructures.

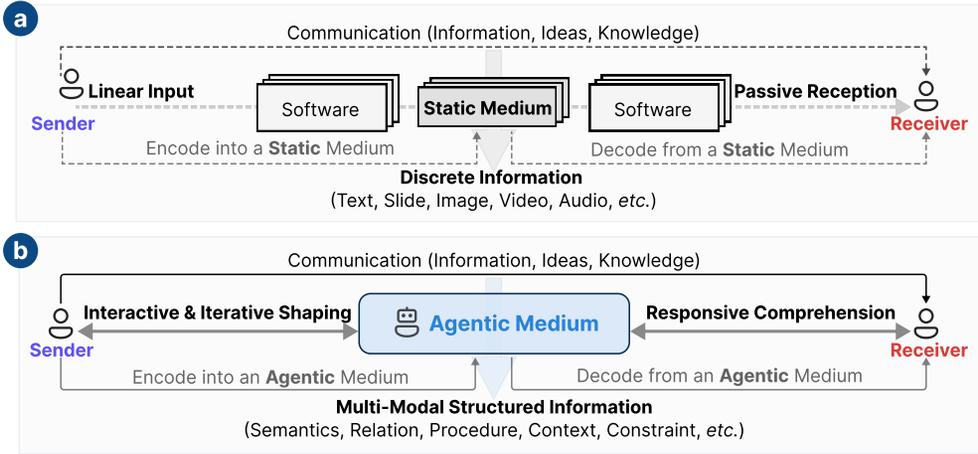


Fig. 2. Comparison between traditional static media and agentic media. In the traditional model, communication is mediated through finalized artifacts produced and consumed via separate software tools. In the Agentic Media model, communication unfolds through ongoing interaction with an integrated medium.

2 The Structural Limits of Format-Centric Media

2.1 Historical Roots of Format-Centric Media and Their Cognitive Impact

Modern digital media formats did not emerge spontaneously; they inherited core structural conventions from earlier physical media. Early software systems deliberately replicated familiar metaphors from physical artifacts to make novel technologies intelligible. For example, the earliest electronic spreadsheet, VisiCalc, mirrored traditional accounting ledgers [12], presentation tools such as PowerPoint adopted the layouts of photographic slides [31], and the Portable Document Format (PDF) emulated printed pages, prioritizing layout fidelity over interactivity [70]. Early graphical user interfaces, including Xerox PARC’s *desktop metaphor* and Alan Kay’s Dynabook, similarly drew on physical-world conventions to ease users’ transition into digital environments [49, 50].

These inherited metaphors brought clear benefits: they stabilized expectations about how information should be organized, manipulated, and shared. At the same time, they encoded particular assumptions about communication into media formats, favoring linear structure, fixed boundaries, and finalized representations. As a result, digital formats such as documents, slides, and spreadsheets tend to privilege certain modes of reasoning and collaboration [30, 40, 41, 89]. This historical legacy helps explain why contemporary knowledge work remains strongly organized around discrete formats, even as technologies and practices evolve.

2.2 Computing Infrastructure and the Technical Reinforcement of Formats

While formats shape how information is conceptualized and shared, their persistence is infrastructural rather than merely cultural. Modern computing hard-codes format-centric logic into its stack: files as atomic units, applications as format-bound editors, operating systems as file coordinators, and devices as display endpoints [5, 32, 49, 93]. Together, these layers establish a pipeline optimized for producing, storing, and distributing finalized artifacts [51, 78, 92, 94].

Within this pipeline, authors produce content in format-specific software, serialize it into discrete files, and distribute them to recipients who primarily view, annotate, or summarize the results [4, 82, 84]. Files become boundary objects detached from their creation contexts, while application boundaries and storage hierarchies further limit recomposition and continuity [93]. The

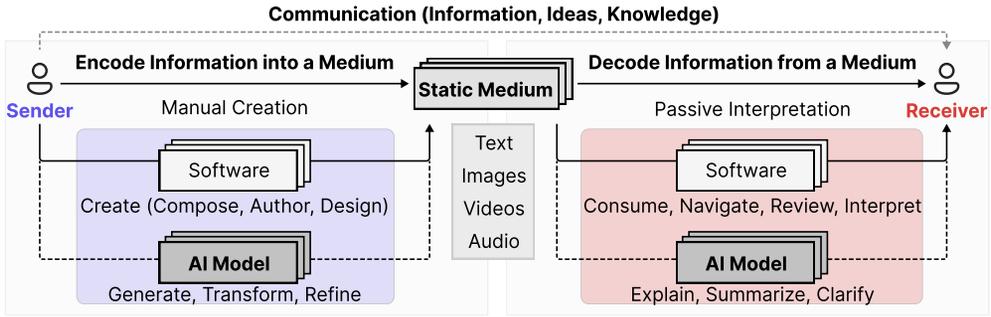


Fig. 3. An overview of contemporary communication pipelines mediated by software and AI. Ideas are encoded into media artifacts through software tools, increasingly with AI assistance, and interpreted by recipients through further transformation, summarization, or explanation. While AI accelerates these transformations, the overall organization of communication remains centered on discrete media artifacts.

sender–receiver model remains dominant, privileging finalized outputs over ongoing processes, which is an inheritance from earlier media paradigms rooted in one-way transmission [87]. Even cloud-native environments largely retain this logic, treating content as static assets to be exported, shared, and archived rather than as evolving communicative processes [18, 51].

Despite technological advances, infrastructures remain organized around publishing rather than communication (Figure 3). Knowledge work becomes locked into rigid formats: workflows bend to templates, artifacts detach from context, and original intent is preserved only as implicit traces [57, 66, 92]. Collaboration is shaped accordingly, collapsing into linear exchanges of files, updates, and summaries that fragment collective understanding [35, 37, 72]. In this setting, productivity tools primarily mediate translation between representations, channeling evolving cognition into transmissible outputs rather than supporting sustained coordination and co-creation [24, 36, 79]. The costs of this organization are tangible. Prior work suggests that knowledge workers spend substantial time searching for and integrating fragmented information across systems [22]. This fragmentation reflects a form of distributed cognition [47], in which understanding is dispersed across people and artifacts, constraining opportunities for reflection-in-action [81].

Recent digital workspace tools have begun to soften the visibility of traditional files and formats by supporting block-based content, conversational streams, and collaborative canvases. While these systems offer more fluid local interactions, they largely operate atop the same underlying infrastructures and do not fundamentally reconfigure artifact-level coordination or responsibility. Moving beyond such increments, Agentic Media targets the structural foundations of digital media themselves, rethinking how communication is coordinated and sustained over time.

2.3 Format-Centric Bias in AI Applications

Recent developments in artificial intelligence reveal a clear pattern: AI tools are increasingly organized around predefined professional roles and established media formats [75]. General-purpose systems routinely produce standardized documents, slide decks, and spreadsheets, while domain-specific agents in areas such as law, healthcare, and finance generate familiar artifacts including contracts, reports, and analyses [60]. Rather than rethinking how communication is structured, these tools reinforce existing divisions of labor and format-bound modes of work.

Beyond content generation, AI is also used to produce software artifacts themselves, including source code and executable programs. Yet even at this level, automation remains embedded within the same communication topology: code is generated, serialized, and exchanged as discrete files,

and coordination continues to revolve around artifact-level handoffs. As a result, the scale and speed of production increase, while the underlying organization of communication remains unchanged. Despite these advances, both levels of automation remain embedded in what we refer to as a traditional *communication topology*: a structure in which people coordinate work primarily by exchanging or co-editing static files. Within this topology, AI increases the volume and speed of production without altering the underlying connections. As a result, gains in scale are often mistaken for improvements in communication [42, 58]. A particularly revealing pattern of this format-centric bias is the increasingly common *over-generation and over-compression cycle* in AI-mediated communication. A minimal prompt often triggers the production of extensive artifacts, *e.g.*, long documents, detailed presentations, or comprehensive summaries, which are then subsequently condensed by others into brief abstracts, bullet points, or short briefs for downstream use [63]. While this cycle appears efficient, it repeatedly re-encodes communicative intent as finalized artifacts, causing meaning to drift or fragment across successive stages of expansion and compression [65, 76]. Rather than supporting sustained interpretation or shared understanding, such cycles foreground artifact production while leaving the coordination and preservation of meaning unresolved. Taken together, these developments reveal a consistent pattern: while recent systems make local interactions more fluid, they largely operate within the same format-centric structures and treat communication as a sequence of artifact-level exchanges. As a result, they alleviate surface frictions without altering the underlying organization of communication itself.

Existing efforts across AI-assisted authoring, collaboration, and generative systems have expanded what can be produced and how interactions unfold. Agentic Media builds on the developments by shifting attention from individual tools or interactions to communication itself as an object of computation.

3 Agentic Media: A New Paradigm of Communication

Across communication studies, design, and computing, the term *media* has been used to refer to channels of communication, materials for expression, and computational interfaces. Despite this diversity, a shared assumption runs through most existing frameworks: media are treated as structured artifacts through which communication occurs, while the processes that generate, transform, and interpret those artifacts are handled externally by tools or systems. Media, in this view, primarily function as vessels or containers for completed expressions. In this work, we take this assumption as a point of departure. We approach media not only as representational containers, but as integral components of communicative processes that shape how meaning is formed and sustained over time. From this perspective, communication is not reducible to the exchange of finalized artifacts, but unfolds through ongoing interaction.

This orientation builds on a long-standing insight across media theory, cognitive science, and HCI: media are not neutral carriers of information. External representations actively shape cognition, reasoning, and coordination, functioning as extensions of human thought [52, 64, 71]. When most digital media are embodied as static representations, the burden of communication was placed on the clarity, fidelity, and completeness at the moment of production. Once produced, artifacts were expected to stand on their own, leaving interpretation and contextualization largely to humans.

Recognizing media as participants within communicative processes opens the possibility of rethinking how media are positioned in communication. Media can be understood as processually involved elements situated within the construction, negotiation, and stabilization of meaning over time. This reframing motivates the proposal of *Agentic Media* as a new paradigm for understanding digital communication. From this perspective, Agentic Media represent a shift toward active partnership between humans and computational media. They foreground how communication itself is structured, addressing limitations embedded in conventional static media formats.

3.1 Defining Agentic Media

In this paper, we use the term *Agentic Media* to refer to computational media that participate directly in communication by engaging with users during interaction. Under this framing, Agentic Media are distinguished not by the presence of automation or intelligence, but by their role within communicative processes. Whereas traditional digital media primarily transmit completed expressions, Agentic Media remain involved as communication unfolds, supporting interpretation, clarification, and revision as part of ongoing interaction. Meaning is thus not treated as fixed at the moment of production, but as something that develops through use.

To illustrate this shift in role, we adopt the metaphor of the *agent as a medium*. Here, the term *agentic* does not imply fully autonomous or goal-independent agency. Instead, it refers to the medium's participatory orientation within communication. Agentic Media operate in close alignment with human communicative intent and are scoped to the purposes of mediation, explanation, and sensemaking.

Importantly, Agentic Media should be distinguished from general-purpose autonomous agents designed to act in open environments or pursue independent goals. Their agency is bounded by communicative context and intent: they function as intermediaries within communication, rather than as actors operating on the world. Across different communicative arrangements, what defines Agentic Media is not the specific configuration of participants, but the fact that the medium itself participates collaboratively in realizing the communicative act. They foreground contextual engagement, structured interactivity, and iterative meaning-making as core features of communication, addressing limitations embedded in conventional static media formats.

3.2 Comparison with Traditional Media

Having clarified what distinguishes Agentic Media conceptually, we next contrast them with traditional media to make these differences concrete. Table 1 outlines how Agentic Media differ from traditional media along several communicative and cognitive dimensions. These dimensions reveal a shift from static information transmission to adaptive, dialogical knowledge systems. Each row represents a foundational change in how meaning is produced, maintained, and evolved through interaction.

Traditional media operate under a transmission model, where authors encode meaning and audiences decode it passively. Agentic Media instead instantiate a procedural communication model, in which interaction becomes the site of meaning-making. Expressive structures are no longer bound by fixed forms but are semantic and adaptive, responding dynamically to user intent and situational context. Information delivery in traditional systems is pre-defined and linear, demanding full cognitive responsibility from the user. Agentic Media redistribute cognitive effort: the system actively scaffolds interpretation, offering clarification and feedback loops that adjust in real time. Finally, while traditional media require manual revision and redistribution to evolve, Agentic Media embody a living update mechanism that evolves during interaction, continuously re-aligning with authorial goals while integrating emergent user perspectives.

As the comparison indicates, Agentic Media reconceptualize media as morphable communicative AI systems. They do not merely present content interactively but sustain a dynamic interpretive process. This enables joint cognition between human and system — a model more suited for complex cognitive activities such as collaborative analysis, decision support, and reflective sensemaking. In this sense, Agentic Media are not a new content type, but a new epistemic infrastructure for communication and thought.

Dimension	Traditional Media	Agentic Media
Communication Model	One-way transmission	Goal-oriented dialogue integrating author intent and user feedback
Expressive Structure	Fixed formats	Semantic, procedural, and adaptive structures
Information Delivery	Static, pre-defined content	Intent-driven responsive interactions supporting runtime adaptation
Consumption Mode	Linear reading, viewing, or watching	Interactive exploration, clarification, and iteration
Cognitive Load	Entirely borne by the user	Jointly supported by system via clarification, scaffolding, and feedback
Evolution Mechanism	Requires manual revision and distribution	Continuously evolves during interactions while preserving authorial intent

Table 1. Comparison between Traditional Media and Agentic Media

4 Experiencing Agentic Media

This section examines how communication is experienced when media participate actively in the communicative process. With Agentic Media, communication is approached as a process in which meaning unfolds over time through sustained engagement. Participants may enter an interaction with different intentions or levels of understanding, and move fluidly between roles such as explaining, questioning, interpreting, and refining meaning as the process evolves. We ground this discussion in familiar communicative situations and examine how these everyday practices are reorganized when the medium itself maintains context, responds to interaction, and supports continuity across exchanges.

4.1 Agentic Media in Use: Rethinking Creation and Consumption

Traditional digital media largely structure communication around one-way reception, where information is packaged into static artifacts and delivered to readers as finalized outputs. In this model, interpretation is assumed to occur after publication, with readers¹ bearing primary responsibility and much of the interpretive burden for resolving ambiguity, filling contextual gaps, and constructing meaning from a fixed representation. Over time, forms of interactivity are introduced to soften this rigid reception model. For example, interactive documents or responsive webpages allow readers to navigate content and reveal additional details, typically based on predefined conditions. Yet such responsiveness is realized through authored structures and predetermined paths, where adaptation operates by selecting among known alternatives. More recent AI-driven tools (such as ChatPDF [17] and NotebookLM [33]) further shift expectations around reception by normalizing questioning, clarification, and re-expression during use. Readers increasingly treat media not as endpoints to be consumed once, but as materials to be queried, transformed, and revisited. However, these practices largely remain layered atop artifact-centered structures, leaving the underlying passive organization of reception fundamentally unchanged.

In Agentic Media, consumption is no longer organized around receiving and interpreting a finalized artifact. It unfolds as an ongoing process in which meaning is progressively shaped through interaction, revision, and continuity across exchanges. Interpretation emerges through successive

¹Here we use *reader* in a broad sense, encompassing audiences across modalities (readers, viewers, listeners, etc.).

engagements with a medium that actively participates in sustaining context and responding to uncertainty. From the reader's perspective, engaging with Agentic Media is marked by several recurring experiential patterns:

- **Iterative exploration.** Understanding develops through cycles of questioning, clarification, and revisiting. Readers return to earlier ideas from new perspectives, with explanations evolving rather than remaining fixed at first encounter.
- **Dialogue shaping.** Rather than following a predetermined narrative path, readers influence how explanations unfold by posing questions, indicating uncertainty, or requesting elaboration. The communicative trajectory is shaped through successive exchanges instead of one-time interpretation.
- **Contextual continuity.** Meaning accumulates across interactions. The medium carries forward contextual information from prior engagement, allowing interpretation to build cumulatively rather than resetting with each encounter.

Taken together, these patterns reframe consumption as participation in an ongoing communicative process. What distinguishes Agentic Media is not greater interactivity alone, but the reorganization of reception around continuity, negotiation, and progressive meaning-making. Readers engage with media that sustain context and respond to uncertainty across exchanges, allowing meaning to be progressively clarified, reformulated, and revisited over time.

Traditional authoring is organized around the production of fixed artifacts that are completed once and then delivered to an audience. Creators design text, visuals, and structure with the expectation that communication largely concludes at the moment of publication. Interpretation is assumed to follow afterward, placing the burden of understanding on readers encountering a finalized representation. Within this model, communicative *context* and *process* are rarely represented explicitly. They are carried through authors' judgment, experience, and anticipation of audience needs. Adaptation tends to take the form of manual duplication or versioning. These adjustments rely on tacit reasoning, making them difficult to externalize, reuse, or computationally support.

With Agentic Media, the logic of authoring shifts from producing static outputs to explicitly designing how communication unfolds over time. Creators model communicative intent, structure how information may be interpreted and revisited, and anticipate how interaction adapts to different user contexts. They move beyond the role of content producers toward that of intent designers and interaction planners. They articulate semantic structures, possible trajectories of interaction, and conditions under which explanations, emphases, or representations may change. Authors may also anticipate likely questions or points of confusion, and specify how the medium should respond as goals, understanding, or context evolve.

Agentic Media also unsettle the assumption of *one-time publishing* that underlies most traditional media. Different from treating publication as a terminal act, Agentic Media allow artifacts to be instantiated and adapted over time. Authorial intent can be carried forward through feedback, reuse, and revision. Under these conditions, authorship becomes distributed and incremental. Media objects are no longer finalized products but evolving sites of communication which we refer to as *living media objects*.

At the same time, Agentic Media do not position creation in opposition to traditional authoring practices. Static, carefully curated artifacts remain valuable, particularly when authors seek control, stability, or precision. The same communicative intent may be instantiated with varying degrees of adaptivity. What changes is the scope of authoring itself: Agentic Media offer a broader design space in which authors can decide when content remains fixed, when it may adapt, and how interaction unfolds over time. The expanded possibilities of Agentic Media also do not necessarily imply greater demands on authors. Although authoring increasingly involves articulating intent, structure, and

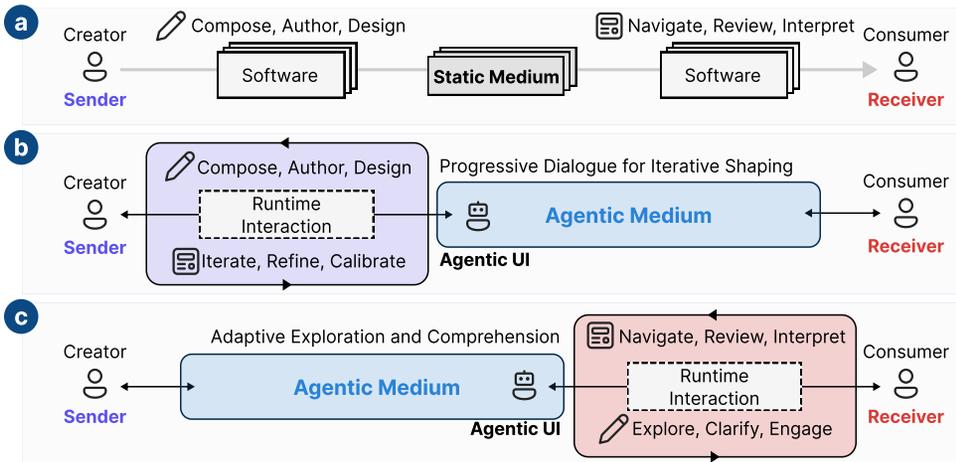


Fig. 4. In the traditional model, creators and consumers engage separately with static media through manual encoding and decoding processes, resulting in fragmented and often inefficient communication. In contrast, the agentic media model integrates creators and consumers into dynamic interactions with an interactive medium. This medium collaboratively assists creators in dynamically encoding content, enabling more precise and semantically accurate expressions. It also supports consumers in achieving a more accurate and efficient understanding, thereby reducing cognitive load and enhancing mutual clarity and communication efficiency.

possible trajectories of interaction, these considerations need not be managed manually or all at once. Computational support may help surface, track, and organize communicative context and interaction histories, allowing authors to work with richer expressive possibilities without being constrained to linear production workflows.

The roles of readers and authors are also reshaped. As interaction becomes responsive, readers are no longer limited to interpreting a fixed artifact. Through questioning, clarification, and revision, they can directly influence how content is reformulated, reorganized, or revisited over time. At the same time, authors gain visibility into how their communicative intent is received, misunderstood, or extended in use. This ongoing feedback loop softens conventional separations between authorship and readership. Contribution and interpretation are no longer strictly separated by publication boundaries, but unfold through interaction. These dynamics also reshape how references and connections between content units function. Instead of static citations or fixed links that serve only as pointers, references can be treated as contextual relationships. Relationships among ideas may be surfaced, elaborated, or contextualized during interaction, reflecting how content is actually used and understood over time.

4.2 Illustrative Scenarios

We describe illustrative scenarios to make the interactional shifts discussed above perceptible. Each scenario highlights a class of problems that conventional media formats handle poorly, and illustrates how these problems are reconfigured when authoring, reception, and memory are treated as ongoing, interactive processes in agentic media.

Problem 1: Sustaining Explanation Through Interaction. Conventional media formats treat explanation as an artifact: authors encode meaning in advance, and readers interpret it after the fact. When misunderstandings arise, they are addressed outside the medium through comments, revisions, or separate exchanges, leaving interpretation structurally decoupled from content. As a

result, understanding is repeatedly restarted. In customer support, this separation is evident in the use of static manuals or FAQ pages designed to deliver fixed information at scale. Once published, these materials cannot evolve with users' questions or contexts. When explanations fail, users must either re-interpret the same content or move the interaction elsewhere, restarting the process rather than refining understanding through interaction.

In agentic media, this separation between explanation and interpretation is reconfigured. Explanations are no longer treated as completed artifacts that must be prepared in advance and repeatedly reinterpreted by users. They remain open to revision through interaction, allowing questions, clarifications, and prior exchanges to directly shape how content is reformulated over time. As a result, understanding can be progressively refined within the medium.

Problem 2: Maintaining Continuity Across Time. Many communicative settings require understanding to unfold across multiple encounters. Conventional educational media such as textbooks, slide decks, or video lectures handle this poorly: when learners return to a concept, they encounter the same representations regardless of prior interaction. Misunderstandings are not carried forward, and progress depends on learners reconstructing context on their own. As a result, learning proceeds as a sequence of disconnected attempts rather than as a continuous trajectory.

A student revisits a difficult concept across several study sessions. Each time, the student returns with a slightly different question or partial understanding shaped by prior attempts. In conventional educational media, this history is invisible to the medium: the same textbook chapter or recorded lecture is encountered again, offering identical explanations regardless of earlier confusion. The student must repeatedly reconstruct context and re-establish understanding, treating each encounter as a fresh start.

With agentic media, this pattern is reconfigured. When the student returns, prior interactions remain available to shape how the material is reintroduced. Explanations are reframed in light of earlier misunderstandings, examples shift to address unresolved questions, and emphasis changes as understanding evolves. Continuity is no longer maintained solely through the student's own memory, but becomes a property of the interaction itself, allowing learning to proceed as a sustained communicative process. This shift extends beyond learners: instructional intent and explanation are no longer fully fixed at the moment of authoring, but remain open to revision through ongoing interaction. This reconfiguration should not be read as another form of teacher agent, but as a structural shift in how information and interaction are organized.

5 Three Layers of Agentic Media

Agentic Media requires new structures that can engage users, manage meaning as an ongoing process, and preserve continuity over time. We introduce three layers to address how these demands are organized within a medium, offering structural perspectives for reasoning about where and how agentic behavior manifests in communicative practice. Broadly, these layers attend to how meaning is surfaced in interaction, coordinated across communicative processes, and carried forward.

The Communicative Interface. At the most visible level, Agentic Media begin where communication meets its expressive surface. In traditional systems, user interfaces are designed for fixed scenarios, each layout, flow, and interaction pattern corresponding to a predetermined set of user goals and scenarios. Even with today's so-called generative UIs, systems largely reproduce new combinations of existing components within pre-authored interaction grammars [19, 44, 56, 91]. They remain mechanisms of user control, structures through which users act upon a system to produce outcomes. What they rarely support is the ongoing process itself, the back-and-forth through which communicative partners align intention and understanding.

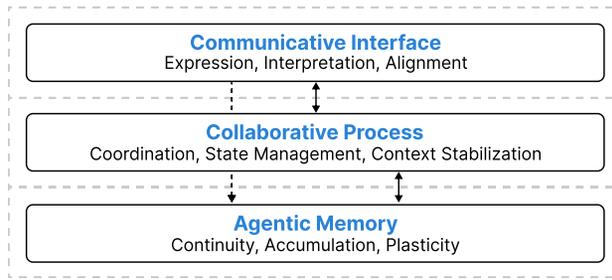


Fig. 5. Agentic Media are organized into three interrelated layers: an interface layer that supports expression, interpretation, and alignment; a process layer that coordinates activity through state management and context stabilization; and a memory layer that carries interaction forward over time.

Within Agentic Media, the interface is reimagined as a co-expressive surface. The Communicative Interface is a shared medium through which both humans and systems articulate, interpret, and adjust meaning. Rather than a static screen or generated layout, it functions as a living space of alignment, where representations shift in response to how communication unfolds. Through multi-modal inputs, adaptive representations, and iterative clarification, such an interface moves beyond enabling actions toward enabling mutual intelligibility. It transforms the surface of interaction from a site of manipulation into a site of co-construction.

This shift dissolves the traditional boundary between interface and content. The interface no longer merely frames interaction or executes commands; it participates in the formation of understanding. In this sense, the design of interaction becomes inseparable from the communication. Media cease to be containers or channels and begin to act as partners in sense-making.

The Collaborative Process. Communication involves many kinds of work: explaining and elaborating, questioning and clarifying, gathering evidence, forming and revising arguments, and adjusting tone or structure for different audiences. These activities rarely occur in a fixed sequence. They overlap and shift as people respond to one another, continually redefining what needs to be said and how it should be understood. To communicate is therefore to manage an evolving process that balances expansion and condensation, maintains coherence across changing perspectives, and sustains mutual understanding over time. From a computational perspective, communication may be understood as a form of process management in organizing collaboration around meaning.

In traditional computing, process management refers to how an operating system schedules execution, allocates resources, and coordinates concurrent operations. Borrowing this idea, the process layer of Agentic Media can be viewed as a collaborative management system for communication. It still involves recognizable tasks such as explaining a point, summarizing progress, or reframing an argument, while extending beyond execution itself. The goal here is to manage participation: aligning intentions, maintaining shared context, and coordinating how communicative actions unfold. The medium may track goals, manage references, detect ambiguity, or suggest rhetorical strategies such as elaboration, comparison, and synthesis. Humans, in turn, tend to undertake the more reflective and higher-level aspects of communication, such as setting priorities, shaping goals, and defining the strategic direction through which meaning evolves. The boundary between human and computational activity remains flexible, and the process layer helps their combined actions stay coherent as contexts evolve.

When communication involves multiple participants, the demands of process management become more adaptive and relational. Several contributors may interact through the same medium

in parallel, creating, reviewing, and responding, while their roles shift as understanding develops. The medium is expected to sustain a shared state of communication, record decisions, and preserve continuity as the interaction evolves.

In a broader view, the process layer sketches a computing model oriented to coordination and interpretation. Its responsibility extends from organizing communicative tasks to governing agency: allocating initiative between people and the medium, deciding when to hand off or escalate, and restoring control when context or stakes change. Therefore, agency is both a capability of the medium and a resource to be managed across participants. By managing initiative, responsibility, and continuity together, the collaborative process layer enables collaboration to scale while keeping goals, roles, and meanings aligned.

The Agentic Memory. In conventional systems, information is stored as predefined data (e.g., files, records, or database entries) that capture interactions as discrete snapshots and finalized outcomes. These structures preserve what has already happened but remain separate from the processes that produced them. They store results rather than continuities. As Agentic Media operate through ongoing interaction, such static forms of memory become insufficient.

A medium that manages communication as a process must also sustain the traces, revisions, and reinterpretations that accumulate within it. Here, memory is no longer an archive of finished states. It becomes a space where knowledge and experience continue to evolve.

Within this space, different forms of memory can coexist: transient buffers for immediate context, episodic traces of interaction, stable conceptual structures, procedural routines, or reflective records of learning and change [90]. Together they enable continuity across time and interaction, allowing the medium to remember not only what was said but also how meaning developed. Accumulated memory turns storage into participation. It carries forward the unfinished aspects of prior exchanges and links them to what follows. In practice, this may appear as persistent records such as annotated dialogues or evolving drafts; semantic representations that abstract recurrent communicative structures into conceptual knowledge; or adaptive behavior, where memory informs how the medium adjusts its tone, modality, or explanatory strategy.

Technically, such memory extends beyond the schema-bound logic of files and databases. Instead of storing predefined structures, it may integrate both structured and adaptive representations: some parts remain explicit and verifiable, while others emerge from patterns formed through learning and interaction. These modes of representation can coexist within the same medium, linking stable knowledge with experience that reorganizes itself over time. The specific mechanisms of such memory are left open for future exploration. What the present discussion of Agentic Media emphasizes is the shift: from storing predefined results toward maintaining an evolving continuity. Therefore, the memory layer supports the plasticity of Agentic Media, enabling information, interpretation, and experience to accumulate and adjust over time, so that the medium can sustain a developing continuity of understanding.

6 Research Opportunities and Challenges

6.1 Why Now: Readiness Across Models, Infrastructure, and Mental Models

Recent advances across AI models, supporting infrastructure, and user expectations jointly make Agentic Media newly plausible as a systematic direction rather than a speculative idea. First, recent generations of large language models have expanded the feasible interaction envelope: long-context processing, real-time multimodal reasoning, and higher-quality on-the-fly synthesis enable media to sustain richer interactive trajectories than earlier systems could support. Second, the maturation of deployment infrastructure lowers the barrier from prototypes to reliable systems. Platforms such as AWS Bedrock [83] and OpenAI Assistant [73] increasingly integrate long-term memory,

function calls, and Retrieval-Augmented Generation (RAG), while open-source frameworks such as LangChain [55] standardize orchestration and deployment pipelines. Together, these reduce engineering friction and make persistent, tool-using, memory-augmented interaction a practical default. Most critically, user mental models are shifting. Rapid adoption signals of AI suggest that interactive, participatory expectations are becoming normalized. Under these conditions, treating communication as an ongoing process becomes thinkable and increasingly demanded.

6.2 Intent as a First-Class Citizen

As analyzed in Section 2, contemporary digital authoring systems implicitly encode authors' communicative intent within rigid formats. For instance, when authors choose a particular slide layout or heading style in PowerPoint or Word, their intent, such as emphasizing a key concept, summarizing complex information, or signaling a transition between topics, is obscured within visual or structural conventions. This implicit treatment of intent has several critical consequences. First, it restricts the capacity of digital systems to dynamically adapt content to varying audience backgrounds, knowledge levels, or interaction histories, thereby forcing content consumers to manually interpret or reconstruct the original communicative purpose. Second, implicit intent reduces content reusability, as secondary authors or agents tasked with remixing or repurposing media artifacts have to reverse-engineer authorial intention from indirect cues. Third, it prevents emerging intelligent agents and adaptive systems from effectively interacting with or dynamically reshaping media content in personalized or contextually adaptive scenarios.

Agentic Media propose treating intent as an explicit, first-class citizen within media creation and interaction. Rather than embedding intent implicitly within fixed formats, authors could explicitly declare communicative goals, anticipated contexts, and desired user interactions, either through structured markup, natural-language annotations, or dynamically inferred intent representations captured during authoring. Such explicit representations of intent, potentially implemented as intent graphs or metadata annotations, could empower intelligent rendering engines and interactive agents to dynamically compile, reorganize, and present content, adapting presentation logic, detail granularity, narrative pathways, and interaction modalities in real-time according to audience characteristics and evolving interaction contexts.

Future research might explore whether authors prefer explicitly declaring intent via structured prompts or annotations, or whether implicit inference mechanisms based on authoring actions could reliably infer intent without disrupting creative flow. Additionally, there is significant scope to investigate computational models and architectures that dynamically interpret and respond to explicitly encoded intent, leveraging interaction paradigms such as adaptive hypermedia, intent-driven dialogue systems, or interactive storytelling [20, 74].

Evaluating intent-centric authoring and compilation systems requires rigorous empirical methods and clearly defined metrics. Prior work on adaptive hypermedia [13], cognitive load theory [86], organizational memory [3], and human-AI interaction guidelines [8] suggests useful criteria. Key evaluation dimensions include content adaptability, reduced cognitive burden, increased rates of content reuse, and improved user satisfaction stemming from enhanced context-awareness and interaction responsiveness.

6.3 From Static Files to Dynamic Communication

At the architectural level, contemporary digital ecosystems are anchored in a file-centric organization, structured around layers of files, file systems, applications, and operating systems. Within the framing developed above, replacing files with databases does not fundamentally alter this representational rigidity. While this design has historically facilitated interoperability and platform standardization, it also reinforces a linear workflow (*creation* → *distribution* → *consumption*) in



Fig. 6. A conceptual view of Communication Architecture in Agentic Media. One way to reason about how communicative intent, semantic structure, interaction, and memory may be coordinated across author and viewer runtimes.

which knowledge is encapsulated in finalized artifacts and meaning must be re-activated through separate acts of human interpretation.

Agentic Media motivates rethinking this organization from the perspective of communication-as-process. Figure 6 provides one conceptual lens for reasoning about where such structure might reside when media participate directly in communication. Central to this lens is the notion of **Communication Architecture**: a computational representation of communicative intent and informational structure that is typically left implicit—carried only in human cognition or informal artifacts such as paper notes or whiteboard sketches [45, 68, 77, 80]. Communication Architecture may encode (or help infer) components such as communicative goals, content organization, narrative structure, and interaction strategies. Externalizing these components into manipulable representations can enable runtime adaptation, and more fluid human-AI collaboration, reducing the burden on users who otherwise have to repeatedly interpret and restructure content across contexts. Importantly, Communication Architecture is not committed to a single representational form. It may be explicit and inspectable, neural and distributed, or hybrid, combining symbolic structure with learned representations [11]. Determining when these representations should be visible, editable, or purely internal remains an open research question spanning AI and cognitive science. A foundational principle, however, is consistent: the medium should computationally assist users in structuring, organizing, and interpreting information in ways that improve communicative efficiency and cognitive usability [8].

Finally, this perspective suggests a bridge between human-facing interaction surfaces and machine-facing internal representations [54, 62]. Rather than treating interfaces as static shells around opaque models, Agentic Media invites a continuous loop in which intent, context, and interaction history are progressively encoded, interpreted, and revised across internal and external layers. In this sense, Communication Architecture functions as a cognitive scaffold for communication, and a design space for exploring how communication-centric computation could be organized.

6.4 Dynamic Co-creation between Creators and Receivers

Compared to traditional “*authored once, then consumed passively*” paradigm, Agentic Media propose a different paradigm of “*dynamic, continuous co-creation*.” Rather than treating media as finalized artifacts, Agentic Media position them as perpetually evolving interactive objects. Creators and consumers interact continuously and symmetrically, jointly influencing content through ongoing feedback loops and iterative co-creation. This continuous co-creation not only enriches the original content through sustained interaction but also extends content lifespan and relevance by making it responsive to the evolving intentions and contexts of diverse user groups. Realizing this vision

demands both conceptual innovation and technical infrastructure. Specifically, it requires embedding flexible, real-time interaction channels directly within media, enabling immediate feedback, annotation, modification, and restructuring. Technically, content should be dynamically compiled and adaptively rendered based on continuous streams of user interaction data and contextual cues, transforming static presentations into living communication processes.

With the media itself serving as a collaborative environment for iterative knowledge building, the boundaries between authorship and readership blur. Future research and design opportunities in this area include developing open-ended content structures capable of accommodating real-time user modifications, building responsive feedback interfaces and interaction channels directly into media frameworks, and engineering technical platforms that support seamless dynamic restructuring and adaptive compilation. Critical evaluation metrics might include measuring content evolution frequency, the depth and quality of user-generated interactions, and improved alignment of media content with evolving user needs and contextual dynamics [6, 8]. Ultimately, the continuous co-creation enabled by Agentic Media turns passive consumption into active participation and positions media as ongoing, collaborative dialogues.

An important aspect of Agentic Media lies in its inherent autonomy, introducing uncertainty beyond the original intent of authors. Specifically, AI-driven mechanisms within Agentic Media may autonomously extend, interpret, or modify content in ways not explicitly envisioned by creators. *Effectively representing and managing the scope of this autonomy is thus a crucial challenge.* Reflecting further, media inherently involves ambiguity in semantic communication, no medium fully captures or transmits an author's exact intention without interpretation. Similarly, Agentic Media embraces this interpretive nature rather than seeking unattainable semantic precision. Its objective is not absolute accuracy, but more effective semantic expression through runtime adaptability and interaction-driven refinement. Ultimately, the efficacy of Agentic Media will depend heavily on thoughtful and precise system design choices, as well as user interaction patterns. Ensuring optimal alignment between AI autonomy and authorial intent remains a central challenge and a promising direction for future research.

6.5 Role Fluidity and Agency

Digital media today typically enforce distinct user roles, separating authors who produce content, readers who consume it, editors who revise it, and reviewers who critique it [38, 57, 94]. While this separation simplifies workflows, it also restricts engagement by confining individuals to narrowly defined, passive interactions. Agentic Media disrupt this rigid model by advocating for **role fluidity**: seamless transitions among traditionally distinct roles. In this paradigm, boundaries between creator and consumer, author and reader, or editor and reviewer are intentionally blurred, allowing users to shift roles as their intentions and contexts evolve. Such fluidity not only enhances user agency, enabling immediate contributions or revisions, but also fosters richer collaboration and collective knowledge generation. This vision goes beyond mixed-initiative interaction [46]: roles in communication are not only negotiated between humans and systems, but also among human participants themselves, blurring the boundaries between author, editor, reviewer, and reader. Achieving role fluidity requires both conceptual and technical advances. Interfaces can encourage smooth transitions from passive consumption to active co-creation, while technical foundations might include fully editable content objects, transparent edit histories, and adaptive feedback mechanisms that suggest role shifts based on user behavior. By enabling role fluidity, Agentic Media transform media consumption into active participation.

6.6 Adaptive Interaction Surfaces

One of the central technical challenges is bridging the gap between the static nature of traditional interface layouts and the inherently dynamic, context-sensitive interactions users increasingly expect. Existing media interfaces are poorly equipped to respond to users' real-time behaviors, rapidly changing intentions, or evolving environmental contexts. Thus, a critical frontier in Agentic Media research lies in developing adaptive interaction surfaces that are capable of real-time, personalized adjustments to user actions, intentions, and contexts.

Recent HCI research has begun to explore this vision of adaptivity [48]. For example, task-driven data models can directly generate user interfaces that change in response to underlying transformations [14]. Chat-based interfaces have been enhanced with adaptive GUI components that surface controls or visualizations on demand, expanding conversational interaction into richer multimodal workflows [39, 44]. Other work has introduced multi-layered textual exploration, where documents reorganize or unfold according to user goals [85]. These efforts highlight an expanding horizon for adaptive interaction, suggesting directions toward dynamic content presentation, conversational augmentation, and reconfigurable graphical components. Delivering such adaptivity at scale, however, requires system-level support. Approaches such as event-driven architectures [43], real-time context processing frameworks [2, 15], and lightweight runtime models to ensure responsiveness and coherence in dynamic interfaces. Key challenges include managing latency so that adaptations feel seamless [16, 69], maintaining state coherence across dynamic updates [88], and executing models efficiently. Looking ahead, opportunities remain in advancing dynamic user modeling [13, 29] and proactive interaction prediction [26, 46]. Establishing evaluation metrics—such as adaptation rate, improvements in user task performance, user satisfaction with responsiveness, and quality of dynamically generated content—will be critical for systematically validating progress.

6.7 Toward an Agentic Media Ecosystem

Current applications are built as isolated, task-specific tools. AI is similarly embedded in narrow domains, such as virtual assistants for daily tasks, specialized design software, or code auto-completion engines. Each of these applications manages its own interaction logic, memory, and context in isolation. As a result, this isolated architecture inevitably leads to fragmented knowledge management and inconsistent user experiences [92]. Users often find their workflow disrupted by having to switch between separate applications, each unaware of interactions or context established elsewhere. Addressing these challenges calls for transitioning toward a more unified and general-purpose Agentic Runtime environment across three aspects: tools, infrastructure, and ecosystems. From the tool view, there needs to be an evolution from conventional, format-bound editors to new authoring platforms explicitly modeling communicative intent and supporting dynamic content generation. From the infrastructure view, digital systems should transition from traditional static file systems toward interactive runtime environments capable of dynamically responding and adapting. From the ecosystem view, it is essential to change from traditional linear broadcast-oriented models toward ecosystems emphasizing ongoing, interactive, and collaborative meaning-making. Achieving these transitions individually is insufficient; only when tools, infrastructure, and ecosystems transform cohesively can Agentic Media move beyond isolated experiments and fully integrate into everyday digital communication. Looking ahead, pursuing this vision raises difficult but exciting research challenges: how to migrate from today's fragmented tools without disrupting existing practices; how to design infrastructures that manage latency, memory, and context coherently at scale; and how to foster ecosystems that align incentives for developers, communities, and industries. Meeting these challenges will be essential to transform digital media into dynamic, interactive spaces that continuously co-create and evolve shared knowledge.

6.8 Governance, Ethics, and Sustainable Systems

Deploying Agentic Media at scale will inevitably surface questions of privacy, transparency, authenticity, and accountability. The very qualities that make these systems powerful, *e.g.*, real-time adaptation, continuous co-creation, and deep personalization, also intensify risks of misuse, algorithmic bias, or erosion of trust. Addressing these challenges requires embedding responsible governance directly into both technical design and ecosystem planning. Emerging *privacy-preserving and transparency-enhancing methods*, such as federated learning, differential privacy, and explainable AI, offer promising avenues. They allow systems to adapt or personalize while reducing exposure of sensitive information and making decision logic more interpretable [7, 21]. Embedding such safeguards into dynamic, large-scale media environments remains an open research frontier, and doing so will be crucial to ensure that Agentic Media develops not only as a technically powerful paradigm but also as one aligned with societal values of trust and accountability.

7 Conceptual Lineage and Related Perspectives

The term media has grown increasingly diffuse. In communication studies, media are seen as vehicles of meaning; in design and HCI, as materials for interaction; and in computing, as information interfaces. This breadth has been historically productive but has also blurred the boundary between media and the tools used to construct them.

In most existing frameworks, authoring tools operate externally to media formats: humans employ tools to create the media artifacts through which communication occurs [27, 61]. For instance, a presentation tool is not itself a medium but a means for producing one. In this paper, we retain this formal distinction to define media as structured containers for communication that carry and shape meaning.

A long-standing body of work across media theory and cognitive science has emphasized that media are not neutral carriers of information, but actively shape cognition, social interaction, and cultural practice. In 1964, Marshall McLuhan famously asserted that “*the medium is the message*,” highlighting that the inherent characteristics of any medium profoundly shape cognitive processes, social interactions, and cultural patterns [64]. Cognitive science further expanded upon this insight by examining how external representations function as extensions of human cognition. Edwin Hutchins introduced the concept of distributed cognition, demonstrating that cognitive processes extend beyond individual minds to include interactions with environmental structures and social contexts [47]. David Kirsh similarly emphasized how external representations such as physical arrangements and visualizations reduce memory load while actively facilitating cognition by transforming abstract problems into spatial and perceptual tasks [52].

Related theories further developed this perspective by examining the role of external artifacts in shaping thinking and action. Andy Clark’s theory of the extended mind argues that cognitive processes depend fundamentally on ongoing interactions with external media [23]. Don Norman likewise stresses that representations shape cognitive outcomes by structuring how people perceive, interpret, and act on information [71]. Within HCI and visualization research, Edward Tufte and Jakob Nielsen have critiqued rigid digital formats for constraining cognitive flexibility. Tufte criticizes linear, slide-based presentations for oversimplifying complex ideas and restricting cognitive depth [89], while Nielsen emphasizes that static, non-interactive document formats such as PDF discourage exploratory reading and limit dynamic collaboration [70].

Parallel to these cognitive perspectives, early visions of computational media sought to move beyond linear textual formats by treating media as active partners in thought. Douglas Engelbart proposed computational systems designed to augment human intellect through non-linear, interactive environments capable of supporting collaboration and knowledge management [28].

Ted Nelson's hypertext and hypermedia concepts similarly emphasized associative, networked structures that better reflect human thought processes [67]. Alan Kay's notion of the "active essay" further expanded this vision by proposing dynamic content that could simultaneously function as explanatory medium and exploratory tool [50]. Building on these ideas, Aarseth's concept of Cybertext reframed media as systems with intrinsic feedback loops, where meaning emerges through traversal rather than passive reception [1].

Despite their conceptual influence, these early visions remained largely theoretical, constrained by the technological limits of their time. Media could invite interaction, but they could not themselves participate in it. Advances in generative artificial intelligence introduce a different technological substrate, extending computation from processing information to participating in communication. Within this context, Agentic Media can be understood as synthesizing insights from distributed cognition, hypertext theory, and generative AI, reframing media from static end-products into adaptive conversations that negotiate meaning in use. Rather than placing the burden solely on representation, Agentic Media foreground the dynamic interplay of sender, receiver, and medium, uniting representation and transmission as inseparable dimensions of the communicative process.

8 Scope and Limitations

This paper offers a conceptual framework. Its illustrative cases are intended to provoke reflection rather than to demonstrate technical feasibility. Questions of autonomy, agency, and internal mechanism design are intentionally left open and are not addressed in this paper. This scope entails important limitations. Technically, open challenges remain in achieving semantic alignment, ensuring interaction coherence, and handling ambiguity at scale. Methodologically, future empirical studies are needed to examine adoption trajectories, long-term use, and potential failure modes. Socially, agentic communication raises questions of authorship, accountability, and governance that extend beyond interface design. Addressing these limitations will require sustained collaboration across HCI, systems design, and the social sciences.

9 Conclusion

This paper set out to examine a largely unspoken assumption underlying contemporary digital media: that communication can be adequately modeled as the production and transfer of finalized representations. By tracing how format-centric infrastructures once stabilized coordination, and how this stabilization erodes as AI lowers the cost of generation and re-expression, we argued that the central challenge of knowledge work has shifted. The difficulty now lies less in producing content than in sustaining communication as meaning evolves across interaction, context, and time. We introduced *Agentic Media* as a conceptual response to this shift. Agentic Media frame communication itself as an active, operable process, one in which intent, interaction, and history can be explicitly represented and continuously negotiated. Agentic Media are not presented as a finalized solution or a closed taxonomy, but as a way of naming a structural reorientation in how media participate in meaning-making. Agentic Media do not simply extend existing tools or interfaces. It is a different way of reasoning about media, computation, and collaboration, one that foregrounds communication as a first-class computation object. By articulating this perspective, we hope to make visible a set of assumptions that have quietly shaped digital work, and to offer a shared conceptual anchor for thinking about how communication-centered media might be understood, critiqued, and explored. The history of media is a history of transformation: from papyrus to paper, from pixels to prompts, each shift has reshaped what we express and how we think together. In this context, Agentic Media mark a possible next step, one that reframes the relationship between humans and media around ongoing communication, where authorship becomes less singular, communication more adaptive, and knowledge more explicitly processual.

References

- [1] Espen J Aarseth. 1997. *Cybertext: Perspectives on ergodic literature*. JHU Press.
- [2] Gregory D Abowd, Anind K Dey, Peter J Brown, Nigel Davies, Mark Smith, and Pete Steggles. 1999. Towards a better understanding of context and context-awareness. In *International symposium on handheld and ubiquitous computing*. Springer, 304–307.
- [3] Mark S. Ackerman and Christine Halverson. 1998. Considering an Organization’s Memory. In *Proceedings of the 1998 ACM Conference on Computer Supported Cooperative Work (CSCW ’98)*. 39–48. doi:10.1145/289444.289461
- [4] Mark S. Ackerman and Christine Halverson. 2004. Organizational Memory as Objects, Processes, and Trajectories: An Examination of Organizational Memory in Use. *Computer Supported Cooperative Work (CSCW)* 13, 2 (2004), 155–189. doi:10.1023/B:COSU.0000045805.77534.2a
- [5] Eytan Adar, David R. Karger, and Lynn Andrea Stein. 1999. Haystack: Per-User Information Environments. In *Proceedings of the 1999 ACM CIKM International Conference on Information and Knowledge Management*. 413–422. doi:10.1145/319950.323231
- [6] Loulwah AlSumait, Daniel Barbará, and Carlotta Domeniconi. 2008. On-line lda: Adaptive topic models for mining text streams with applications to topic detection and tracking. In *2008 eighth IEEE international conference on data mining*. IEEE, 3–12.
- [7] Y. I. Alzoubi. 2025. Differential privacy and artificial intelligence: potentials and challenges. *EURASIP Journal on Information Security* (2025).
- [8] Saleema Amershi, Dan Weld, Mihaela Vorvoreanu, Adam Fourney, Besmira Nushi, Penny Collisson, Jina Suh, Shamsi T. Iqbal, Paul N. Bennett, Kori Inkpen, Jaime Teevan, Ruth Kikin-Gil, and Eric Horvitz. 2019. Guidelines for Human-AI Interaction. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems (CHI ’19)*. ACM, Article 3, 13 pages. doi:10.1145/3290605.3300233
- [9] Alan F Blackwell. 2006. The reification of metaphor as a design tool. *ACM Transactions on Computer-Human Interaction (TOCHI)* 13, 4 (2006), 490–530.
- [10] Verónica Bolón-Canedo, Laura Morán-Fernández, Brais Cancela, and Amparo Alonso-Betanzos. 2024. A review of green artificial intelligence: Towards a more sustainable future. *Neurocomputing* 599 (2024), 128096.
- [11] Ronald Brachman and Hector Levesque. 2004. *Knowledge representation and reasoning*. Elsevier.
- [12] Dan Bricklin. 2007. Why Didn’t We Patent the Spreadsheet? Were We Stupid? <http://www.bricklin.com/patenting.htm>. Accessed July 29, 2025.
- [13] Peter Brusilovsky. 2001. Adaptive hypermedia. *User Modeling and User-Adapted Interaction* 11, 1-2 (2001), 87–110. doi:10.1023/A:1011143116306
- [14] Yining Cao, Peiling Jiang, and Haijun Xia. 2025. Generative and Malleable User Interfaces with Generative and Evolving Task-Driven Data Model. In *Proceedings of the 2025 CHI Conference on Human Factors in Computing Systems*. 1–20.
- [15] Paris Carbone, Asterios Katsifodimos, Stephan Ewen, Volker Markl, Seif Haridi, and Kostas Tzoumas. 2015. Apache Flink: Stream and Batch Processing in a Single Engine. *Bulletin of the IEEE Computer Society Technical Committee on Data Engineering* 36, 4 (2015), 28–38.
- [16] Stuart K. Card, Thomas P. Moran, and Allen Newell. 1983. *The Psychology of Human-Computer Interaction*. Lawrence Erlbaum Associates.
- [17] ChatPDF. 2025. ChatPDF: Chat with any PDF. <https://www.chatpdf.com> Accessed July 29, 2025.
- [18] Souti Chattopadhyay, Ishita Prasad, Austin Z. Henley, Anita Sarma, and Titus Barik. 2020. What’s Wrong with Computational Notebooks? Pain Points, Needs, and Design Opportunities. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*. 1–12. doi:10.1145/3313831.3376729
- [19] Jiaqi Chen, Yanzhe Zhang, Yutong Zhang, Yijia Shao, and Diyi Yang. 2025. Generative interfaces for language models. *arXiv preprint arXiv:2508.19227* (2025).
- [20] Weihao Chen, Chun Yu, Huadong Wang, Zheng Wang, Lichen Yang, Yukun Wang, Weinan Shi, and Yuanchun Shi. 2023. From gap to synergy: Enhancing contextual understanding through human-machine collaboration in personalized systems. In *Proceedings of the 36th Annual ACM Symposium on User Interface Software and Technology*. 1–15.
- [21] Ben Chester Cheong. 2024. Transparency and accountability in AI systems. *Frontiers in Human Dynamics* (2024).
- [22] Michael Chui, James Manyika, Jacques Bughin, Richard Dobbs, Charles Roxburgh, Hugo Sarrazin, Geoffrey Sands, and Magdalena Westergren. 2012. *The Social Economy: Unlocking Value and Productivity through Social Technologies*.
- [23] Andy Clark. 2008. *Supersizing the Mind: Embodiment, Action, and Cognitive Extension*. Oxford University Press, Oxford.
- [24] Samuel Rhys Cox, Helena Bøjer Djernæs, and Niels van Berkel. 2025. Beyond Productivity: Rethinking the Impact of Creativity Support Tools. *arXiv preprint arXiv:2505.01601* (2025). <https://arxiv.org/abs/2505.01601>
- [25] Laura Dabbish, Colleen Stuart, Jason Tsay, and Jim Herbsleb. 2012. Social coding in GitHub: transparency and collaboration in an open software repository. In *Proceedings of the ACM 2012 conference on computer supported cooperative work*. 1277–1286.

- [26] Hai Dang, Sven Goller, Florian Lehmann, and Daniel Buschek. 2023. Choice Over Control: How Users Write with Large Language Models using Diegetic and Non-Diegetic Prompting. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems (CHI '23)*. ACM, New York, NY, USA, Article 408, 17 pages. doi:10.1145/3544548.3580969
- [27] Lars Elleström. 2010. The modalities of media: A model for understanding intermedial relations. In *Media borders, multimodality and intermediality*. Springer, 11–48.
- [28] Douglas C. Engelbart. 1962. *Augmenting Human Intellect: A Conceptual Framework*. Technical Report AFOSR-3223. Stanford Research Institute, Menlo Park, CA. Summary report prepared for the Air Force Office of Scientific Research.
- [29] Gerhard Fischer. 2001. User modeling in human–computer interaction. *User modeling and user-adapted interaction* 11, 1 (2001), 65–86.
- [30] Jonas Frich, Lindsay MacDonald Vermeulen, Christian Remy, Michael Mose Biskjaer, and Peter Dalsgaard. 2019. Mapping the landscape of creativity support tools in HCI. In *Proceedings of the 2019 CHI conference on human factors in computing systems*. 1–18.
- [31] Robert Gaskins. 2012. *Sweating Bullets: Notes about Inventing PowerPoint*. Vinland Books, San Francisco, CA and London, UK.
- [32] David K. Gifford, Pierre Jouvelot, Mark A. Sheldon, and Jr. O’Toole, James W. 1991. Semantic File Systems. In *Proceedings of the Thirteenth ACM Symposium on Operating Systems Principles*. 16–25. doi:10.1145/121132.121138
- [33] Google. 2023. AI Products and Documentation: NotebookLM, Gemini 1.5 Pro, and Workspace AI. <https://ai.google/>
- [34] Tanya Goyal and Greg Durrett. 2021. Annotating and Modeling Fine-grained Factuality in Summarization. *arXiv preprint arXiv:2104.04302* (2021). <https://arxiv.org/abs/2104.04302>
- [35] Jonathan Grudin. 1994. Groupware and Social Dynamics: Eight Challenges for Developers. In *Communications of the ACM*, Vol. 37. 92–105. doi:10.1145/175222.175230
- [36] Mengyao Guo, Jinda Han, Ze Gao, Yuan Zhuang, and Xingting Wu. 2025. Human and Machine as Seen at the Co-Creation Age: A Co-Word Analysis in Human Machine Co-creation (2014-2024). *arXiv preprint arXiv:2505.14363* (2025). <https://arxiv.org/abs/2505.14363>
- [37] C. Gómez and collaborators. 2025. A taxonomy of interaction patterns in AI-assisted decision making. *Frontiers in Computer Science* (2025). doi:10.3389/fcomp.2024.1521066
- [38] Yuanning Han, Ziyi Qiu, Jiale Cheng, and Ray LC. 2024. When Teams Embrace AI: Human Collaboration Strategies in Generative Prompting in a Creative Design Task. In *Proceedings of the 2024 CHI Conference on Human Factors in Computing Systems*.
- [39] Marti Hearst and Melanie Tory. 2019. Would you like a chart with that? Incorporating visualizations into conversational interfaces. In *2019 IEEE Visualization Conference (VIS)*. IEEE, 1–5.
- [40] Luke Hespanhol. 2023. Human-computer intra-action: a relational approach to digital media and technologies. *Frontiers in Computer Science* (2023). doi:10.3389/fcomp.2023.1083800
- [41] Otmar Hilliges, Lucia Terrenghi, Sebastian Boring, David Kim, Hendrik Richter, and Andreas Butz. 2007. Designing for collaborative creative problem solving. In *Proceedings of the 6th ACM SIGCHI conference on Creativity & cognition*. 137–146.
- [42] Jess Hohenstein, René F. Kizilcec, Dominic DiFranzo, Zhila Aghajari, Hannah Mieczkowski, Karen Levy, Mor Naaman, Jeffrey Hancock, and Malte F. Jung. 2023. Artificial intelligence in communication impacts language and social relationships. *Scientific Reports* 13, 5487 (2023). doi:10.1038/s41598-023-30938-9
- [43] Gregor Hohpe and Bobby Woolf. 2004. *Enterprise Integration Patterns: Designing, Building, and Deploying Messaging Solutions*. Addison-Wesley.
- [44] Nobukatsu Hojo, Kazutoshi Shinoda, Yoshihiro Yamazaki, Keita Suzuki, Hiroaki Sugiyama, Kyosuke Nishida, and Kuniko Saito. 2025. GenerativeGUI: Dynamic GUI Generation Leveraging LLMs for Enhanced User Interaction on Chat Interfaces. In *Proceedings of the Extended Abstracts of the CHI Conference on Human Factors in Computing Systems*. 1–9.
- [45] James Hollan, Edwin Hutchins, and David Kirsh. 2000. Distributed cognition: toward a new foundation for human-computer interaction research. *ACM Transactions on Computer-Human Interaction (TOCHI)* 7, 2 (2000), 174–196.
- [46] Eric Horvitz. 1999. Principles of Mixed-Initiative User Interfaces. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. ACM, Pittsburgh, PA, USA, 159–166.
- [47] Edwin Hutchins. 1995. *Cognition in the Wild*. MIT Press, Cambridge, MA.
- [48] Anthony Jameson. 2003. Adaptive Interfaces and Agents. In *The Human-Computer Interaction Handbook: Fundamentals, Evolving Technologies and Emerging Applications*, Julie A. Jacko and Andrew Sears (Eds.). Lawrence Erlbaum Associates, Mahwah, NJ, USA, 305–330.
- [49] Jeff Johnson, Teresa L. Roberts, William Verplank, David C. Smith, Charles H. Irby, Marian Beard, and Kevin Mackey. 1989. The Xerox Star: A Retrospective. *Computer* 22, 9 (1989), 11–26. doi:10.1109/2.35601
- [50] Alan C. Kay and Adele Goldberg. 1977. Personal Dynamic Media. *Computer* 10, 3 (1977), 31–41.

- [51] Mary Beth Kery, Marissa Radensky, Mahima Arya, Bonnie E. John, and Brad A. Myers. 2018. The Story in the Notebook: Exploratory Data Science Using a Literate Programming Tool. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*. 1–11. doi:10.1145/3173574.3173748
- [52] David Kirsh. 2010. Thinking with external representations. *AI & Society* 25, 4 (2010), 441–454.
- [53] Ned Kock. 2005. Media richness or media naturalness? The evolution of our biological communication apparatus and its influence on our behavior toward e-communication tools. *IEEE transactions on professional communication* 48, 2 (2005), 117–130.
- [54] Brenden M. Lake, Tomer D. Ullman, Joshua B. Tenenbaum, and Samuel J. Gershman. 2017. Building machines that learn and think like people. *Behavioral and Brain Sciences* 40 (2017), 1–72. doi:10.1017/S0140525X16001837
- [55] LangChain. 2024. LangChain v0.2 Documentation. <https://docs.langchain.com/docs/v0.2>. Accessed August 1, 2025.
- [56] Hugo Laurençon, Léo Tronchon, and Victor Sanh. 2024. Unlocking the conversion of web screenshots into html code with the websight dataset. *arXiv preprint arXiv:2403.09029* (2024).
- [57] Haotian Li, Yun Wang, and Huamin Qu. 2024. Where Are We So Far? Understanding Data Storytelling Tools from the Perspective of Human–AI Collaboration. In *Proceedings of the 2024 CHI Conference on Human Factors in Computing Systems (CHI '24)*. 845:1–845:19. doi:10.1145/3613904.3642726
- [58] Zhuoyan Li, Chen Liang, Jing Peng, and Ming Yin. 2024. The Value, Benefits, and Concerns of Generative AI-Powered Assistance in Writing. *arXiv preprint arXiv:2403.12004* (2024). <https://arxiv.org/abs/2403.12004>
- [59] Alexandra Sasha Luccioni, Emma Strubell, and Kate Crawford. 2025. From efficiency gains to rebound effects: The problem of jevons’ paradox in AI’s polarized environmental debate. In *Proceedings of the 2025 ACM Conference on Fairness, Accountability, and Transparency*. 76–88.
- [60] Varun Magesh, Faiz Surani, Matthew Dahl, Mirac Suzgun, Christopher D. Manning, and Daniel E. Ho. 2025. Hallucination-Free? Assessing the Reliability of Leading AI Legal Research Tools. *Journal of Empirical Legal Studies* (2025), 1–27. doi:10.1111/jels.12413
- [61] Lev Manovich. 2002. The language of new media.
- [62] Gary Marcus. 2020. The Next Decade in AI: Four Steps Towards Robust Artificial Intelligence. <https://arxiv.org/abs/2002.06177>.
- [63] Joshua Maynez, Shashi Narayan, Bernd Bohnet, and Ryan McDonald. 2020. On Faithfulness and Factuality in Abstractive Summarization. In *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics (ACL 2020)*. 1906–1919. doi:10.18653/v1/2020.acl-main.173
- [64] Marshall McLuhan. 1994. *Understanding media: The extensions of man*. MIT press.
- [65] Danielle S. McNamara, Micah Watanabe, Linh Huynh, Kathryn S. McCarthy, Laura K. Allen, and Joseph P. Magliano. 2024. Summarizing versus rereading multiple documents. *Contemporary Educational Psychology* 76 (2024), Article 102238. doi:10.1016/j.cedpsych.2023.102238
- [66] Michael Muller and Justin D. Weisz. 2022. Extending a Human–AI Collaboration Framework with Dynamism and Sociality. In *Proceedings of the 2022 Symposium on Human-Computer Interaction for Work (CHIWORK '22)*. 111:1–111:18. doi:10.1145/3533406.3533407
- [67] Theodor H. Nelson. 1992. *Literary Machines* (93.1 ed.). Mindful Press, Sausalito, CA.
- [68] John F Nestojko, Jason R Finley, and Henry L Roediger III. 2013. Extending cognition to external agents. *Psychological Inquiry* 24, 4 (2013), 321–325.
- [69] Jakob Nielsen. 1993. *Usability Engineering*. Morgan Kaufmann.
- [70] Jakob Nielsen. 2003. PDF: Unfit for Human Consumption. Alertbox, Nielsen Norman Group.
- [71] Donald A. Norman. 1993. *Things That Make Us Smart: Defending Human Attributes in the Age of the Machine*. Addison-Wesley, Reading, MA.
- [72] Gary M. Olson and Judith S. Olson. 2000. Distance Matters. *Human-Computer Interaction* 15, 2 (2000), 139–178. doi:10.1207/S15327051HCI1523_4
- [73] OpenAI. 2021–2025. OpenAI Research and Product Documentation. <https://openai.com/>.
- [74] Cheng Qian, Bingxiang He, Zhong Zhuang, Jia Deng, Yujia Qin, Xin Cong, Zhong Zhang, Jie Zhou, Yankai Lin, Zhiyuan Liu, et al. 2024. Tell me more! towards implicit user intention understanding of language model driven agents. *arXiv preprint arXiv:2402.09205* (2024).
- [75] Mubashar Raza, Zarmina Jahangir, Muhammad Bilal Riaz, Muhammad Jasim Saeed, and Muhammad Awais Sattar. 2025. Industrial applications of large language models. *Scientific Reports* 15 (2025), Article 13755. doi:10.1038/s41598-025-98483-1
- [76] Manoel Horta Ribeiro, Kristina Gligorić, and Robert West. 2019. Message Distortion in Information Cascades. In *Proceedings of The Web Conference 2019 (WWW '19)*. 681–692. doi:10.1145/3308558.3313531
- [77] Evan F Risko and Sam J Gilbert. 2016. Cognitive offloading. *Trends in cognitive sciences* 20, 9 (2016), 676–688.
- [78] Adam Rule, Aurélien Tabard, and James D. Hollan. 2018. Exploration and Explanation in Computational Notebooks. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*. 1–12. doi:10.1145/3173574.3173606

- [79] Advait Sarkar. 2023. Enough With “Human-AI Collaboration”. In *Extended Abstracts of the 2023 CHI Conference on Human Factors in Computing Systems (CHI EA '23)*. 415:1–415:8. doi:10.1145/3544549.3582735
- [80] Mike Scaife and Yvonne Rogers. 1996. External cognition: how do graphical representations work? *International journal of human-computer studies* 45, 2 (1996), 185–213.
- [81] Donald A. Schön. 1983. *The Reflective Practitioner: How Professionals Think in Action*. Basic Books, New York.
- [82] Wilbur Schramm. 1954. *The Process and Effects of Mass Communication*. University of Illinois Press, Urbana, IL.
- [83] Amazon Web Services. 2024. Amazon Bedrock: Agents, Memory, and Retrieval-Augmented. <https://aws.amazon.com/bedrock>.
- [84] Claude E Shannon. 1948. A mathematical theory of communication. *The Bell system technical journal* 27, 3 (1948), 379–423.
- [85] Sangho Suh, Bryan Min, Srishti Palani, and Haijun Xia. 2023. Sensecape: Enabling multilevel exploration and sensemaking with large language models. In *Proceedings of the 36th annual ACM symposium on user interface software and technology*. 1–18.
- [86] John Sweller. 1988. Cognitive load during problem solving: Effects on learning. *Cognitive Science* 12, 2 (1988), 257–285. doi:10.1207/s15516709cog1202_4
- [87] Jaime Teevan, Christine Alvarado, Mark S. Ackerman, and David R. Karger. 2004. The Perfect Search Engine Is Not Enough: A Study of Orienteering Behavior in Directed Search. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. 415–422. doi:10.1145/985692.985745
- [88] Douglas B Terry, Alan J Demers, Karin Petersen, Mike J Spreitzer, Marvin M Theimer, and Brent B Welch. 1994. Session guarantees for weakly consistent replicated data. In *Proceedings of 3rd International Conference on Parallel and Distributed Information Systems*. IEEE, 140–149.
- [89] Edward R. Tufte. 2003. *The Cognitive Style of PowerPoint*. Graphics Press, Cheshire, CT.
- [90] Endel Tulving. 1985. How many memory systems are there? *American psychologist* 40, 4 (1985), 385.
- [91] Priyan Vaithilingam, Elena L Glassman, Jeevana Priya Inala, and Chenglong Wang. 2024. Dynavis: Dynamically synthesized ui widgets for visualization editing. In *Proceedings of the 2024 CHI Conference on Human Factors in Computing Systems*. 1–17.
- [92] Yun Wang and Yan Lu. 2025. Interaction, Process, Infrastructure: A Unified Architecture for Human-Agent Collaboration. arXiv:2506.11718 [cs.HC] <https://arxiv.org/abs/2506.11718>
- [93] Terry Winograd and Fernando Flores. 1986. *Understanding Computers and Cognition: A New Foundation for Design*. Vol. 335. Ablex Publishing, Norwood, NJ.
- [94] Amy X Zhang, Michael Muller, and Dakuo Wang. 2020. How do data science workers collaborate? roles, workflows, and tools. *Proceedings of the ACM on Human-Computer Interaction* 4, CSCW1 (2020), 1–23.