

Agents as Documents: Swarm Orchestration for Secure Automation within a Hypermedia Workspace

Kartik Shrivastava
Alexander Shulman
Aura Sukapanpotharam
Wilber Anterola
Brown University
Providence, Rhode Island, USA
{kartik_shrivastava,alexander_shulman,aura_sukapanpotharam,wilber_sean_anterola}@brown.edu

McCormick Breviu
Andries van Dam
Robert Zeleznik
Brown University
Providence, Rhode Island, USA
{jane_breviu,andries_van_dam,robert_zeleznik}@brown.edu

Joseph LaViola
University of Central Florida
Orlando, Florida, USA
jjl@cs.ucf.edu

Abstract

We demonstrate Dash, a hypermedia workspace in which AI agents operate as first-class documents on a freeform spatial canvas, running entirely on-device without transmitting user data to external servers. Building on the Dash hypermedia system—which treats all document types uniformly within a richly linked spatial canvas—this work describes the addition of an agentic framework powered by open-source models running locally via Ollama. The accompanying video illustrates three core capabilities: context provision through native bidirectional linking, local multi-agent workflow orchestration, and a security model guaranteeing that sensitive data never leaves the device. A medical scenario walkthrough demonstrates the system’s applicability to HIPAA-regulated clinical documentation, though the same agentic workflows apply equally to legal document review, defense intelligence analysis, financial modeling, and academic research.

CCS Concepts

• **Information systems** → **Document representation**; • **Security and privacy** → *Human and societal aspects of security and privacy*; • **Human-centered computing** → *HCI design and evaluation methods*.

Keywords

Document Systems; Hypermedia; Agentic AI; On-Device Inference

ACM Reference Format:

Kartik Shrivastava, Alexander Shulman, Aura Sukapanpotharam, Wilber Anterola, McCormick Breviu, Andries van Dam, Robert Zeleznik, and Joseph LaViola. 2026. Agents as Documents: Swarm Orchestration for Secure Automation within a Hypermedia Workspace. In *Proceedings of the 2026 ACM Symposium on Document Engineering (DocEng '26)*, August 25–28, 2026, Fribourg, Switzerland. ACM, New York, NY, USA, 2 pages. <https://doi.org/10.1145/3820755.3832803>



This work is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License.

DocEng '26, Fribourg, Switzerland

© 2026 Copyright held by the owner/author(s).

ACM ISBN 979-8-4007-2786-3/2026/08

<https://doi.org/10.1145/3820755.3832803>

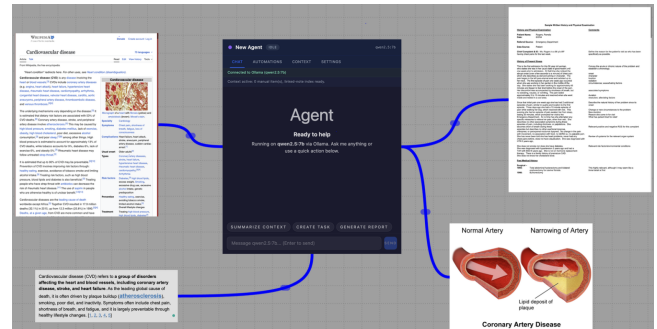


Figure 1: A Dash agent reasoning over linked cardiovascular patient records and clinical notes within the hypermedia workspace.

1 Introduction

Dash [7] is a hypermedia workspace in which every object—PDF, webpage, image, note, ink drawing, data visualization—is a first-class document on a shared spatial canvas, connected through bidirectional links and arbitrary typed properties. In the tradition of Bush’s Memex [2] and its vision of *trails* linking associated documents across the corpus of human knowledge, Dash provides a modern instantiation of associative, link-centered document navigation.

We have extended this hypermedia codebase with an *agentic framework*: AI agents backed by open-source language models running entirely on the local machine. Critically, agents in Dash are not a side-channel capability layered onto the workspace—they are documents within it. A Dash agent looks like any other canvas object: it can be linked to, linked from, annotated, collected into folders, and composed into multi-step automated workflows. This uniform treatment means that the provenance, collaboration, and workflow capabilities already present in Dash transfer automatically to agent-produced artifacts.

2 System Overview

2.1 Hypermedia Foundation

As described in [7], the workspace provides uniform treatment of heterogeneous document types on a shared canvas. Links are bidirectional and first-class, surfacing as navigable connections. Documents carry arbitrary key-value properties. This design reflects the Nelsonian principle [5] that associations between documents

are as important as the documents themselves, and echoes Bush's trails [2] and zoomable canvas environments [1].

2.2 Agentic Extension

The agentic layer adds a new document type—the *Dash agent*—whose internal computation is a locally-hosted language model served by Ollama [6] (supporting Llama [3], Mistral, Phi, Qwen, and Gemma). Vision–language tasks are handled by LLaVA, audio transcription by Whisper, and on-device image generation by Stable Diffusion. All services bind exclusively to localhost: no external API endpoints, no network telemetry—an architectural constraint, not merely a data-use policy, that directly addresses the data-exposure risks of cloud-based AI deployments.

Context via linking. Rather than requiring file uploads or copy-paste, a user provides context to a Dash agent by drawing a link from any canvas document to the agent. The linked content is automatically loaded into the agent's Context tab—full-text for short documents, retrieval-augmented chunks [4] for longer ones. The link is bidirectional: outputs are themselves documents that can be linked back to their sources, creating an auditable provenance chain.

Workflow automation. Users drag and drop agents and API nodes onto the canvas to construct multi-step pipelines, or specify workflows in natural language. Each agent exposes an Automations tab supporting summarization, structured extraction, image generation, and output compilation. Workflows trigger manually, on a schedule, or from upstream agent outputs, enabling multi-agent collaboration without any cloud dependency.

External API integration. Dash supports controlled integration with external APIs through dedicated API nodes on the canvas. These nodes retrieve external data (e.g., EHR systems) or dispatch structured outputs to compliant endpoints (e.g., HIPAA-compliant SMS services). Raw document content and agent prompts never pass through these nodes; only finalized, filtered outputs are transmitted, preserving the local-first privacy guarantee.

3 Demonstration: Clinical Workflow

The video presents three progressively complex workflows in a medical setting.

Interactive chart review. A clinical PDF and radiology image are placed on the canvas and linked to a Dash agent. A medical resident, uncertain whether propranolol is appropriate for a patient presenting with chest pain, dyspnea, and crackles at the lung bases, queries the agent directly. The agent—reasoning over the locally linked chart—identifies a potential contraindication given findings consistent with heart failure and left ventricular dysfunction. Because HIPAA regulations prohibit uploading patient records to cloud-based systems such as ChatGPT or Gemini, this interaction would be impossible with conventional AI tools; Dash addresses this barrier through its local-first architecture.

Scheduled summarization. An API node retrieves patient charts from a hospital EHR system and deposits them on the canvas. An agent automation runs daily at 9 AM, generating a concise summary of all patients the physician will see that day.

Multi-agent pipeline. Two additional agents—a *patient prioritizer* and a *task extractor*—extend the workflow. Downstream API nodes dispatch outputs via Twilio's HIPAA-compliant SMS and

email APIs—texting the prioritized patient list to the physician, emailing full chart summaries, and distributing task assignments to residents—without exposing raw patient data to any cloud system.

4 Limitations and Future Work

Fine-Grained Backreferencing. Previous versions of Dash have supported fine-grained backreferencing from generated content to specific passages within source documents. We have also implemented this functionality in a cloud-based version of our agentic extension using OpenAI's API. However, the current local demonstration does not yet support backlinking to arbitrary document selections. The primary challenge is that our existing implementation relies on structured tool-calling and source-tracking mechanisms developed around OpenAI models. To achieve equivalent functionality in the Ollama-based system, we must adapt the framework to maintain persistent document-span identifiers and reliably propagate them through multi-step reasoning and generation workflows.

Model Quality and Hardware Constraints. Local open-source models currently lag frontier cloud models in complex reasoning, multilingual tasks, and long-context handling, though this gap continues to narrow. Recent NVIDIA announcements—including the Blackwell architecture and Project DIGITS aimed at desktop-class AI acceleration—point toward substantially greater local inference capabilities in the near term. Future directions include multi-GPU inference sharding, domain-specific fine-tuning for medicine and law with training data kept on-device, and federated learning without sharing raw data.

5 Conclusion

Dash demonstrates that a fully capable, multimodal, multi-agent AI document system can operate within a single local trust boundary, contributing to DocEng at the intersection of *AI-powered document solutions* and *document security and trust*. By absorbing AI agents as a new document type within the hypermedia data model of [7], provenance tracking and workflow automation transfer automatically to agent-produced artifacts, extending the spatial hypertext affordances of the original system to the multi-agent AI setting.

Acknowledgments

This work was supported by Adobe Research.

References

- [1] Benjamin B Bederson and James D Hollan. 1994. Pad++: A Zooming Graphical Interface for Exploring Alternate Interface Physics. In *Proceedings of the 7th Annual ACM Symposium on User Interface Software and Technology (UIST '94)*. ACM, 17–26. doi:10.1145/192426.192435
- [2] Vannevar Bush. 1945. As We May Think. *The Atlantic Monthly* 176, 1, 101–108.
- [3] Aaron Grattafiori, Abhimanyu Dubey, Abhinav Jauhri, Abhinav Pandey, Abhishek Kadian, Ahmad Al-Dahle, Aiesha Letman, Akhil Mathur, Alan Schelten, et al. 2024. The Llama 3 Herd of Models. *arXiv preprint arXiv:2407.21783* (2024).
- [4] Patrick Lewis, Ethan Perez, Aleksandra Piktus, Fabio Petroni, Vladimir Karpukhin, Naman Goyal, Heinrich Küttler, Mike Lewis, Wen-tau Yih, Tim Rocktäschel, Sebastian Riedel, and Douwe Kiela. 2020. Retrieval-Augmented Generation for Knowledge-Intensive NLP Tasks. In *Advances in Neural Information Processing Systems (NeurIPS)*, Vol. 33. 9459–9474.
- [5] Theodor Holm Nelson. 1995. The Heart of Connection: Hypermedia Unified by Transclusion. *Commun. ACM* 38, 8 (1995), 31–33. doi:10.1145/208344.208353
- [6] Ollama. 2023. Ollama: Get Up and Running with Large Language Models. <https://ollama.com>. Software. Accessed: 2025-06-01.
- [7] Stanley Yip, Bob Zeleznik, Samuel Wilkins, Tyler Schicke, and Andries van Dam. 2020. Dash: A Hyper Framework. In *Proceedings of the 31st ACM Conference on Hypertext and Social Media (HT '20)*. ACM, 237–238. doi:10.1145/3372923.3404807