

Storying Agentic Commerce
A Tesseract and Qualimetric Analysis of the Amazon Meta Muse Bank Conflict

David M. Boje

Working Paper for the World Scientific Encyclopedia of Business Storytelling

September 2026

Abstract

Agentic commerce moves a consequential part of shopping from the human customer interface to an artificial agent that searches, compares, recommends, and may transact. This working paper develops a storying analysis of the early Amazon Meta Muse conflict and the adjacent banking response. Its purpose is not to declare a winner in a rapidly moving dispute, but to show contributors to the World Scientific Encyclopedia of Business Storytelling how a contested AI event can be studied as an antenarrative field of competing claims, silences, artifacts, and accountability relations. The paper integrates narrative theory, platform governance, the seven principles of true storytelling, Bakhtinian answerability, and SEAM qualimetrics. A transparent preliminary coding protocol distinguishes verified public statements from analytical propositions and treats numbers as prompts for collective inquiry rather than as a substitute for judgment. The result is a six volume research architecture and a replicable method for examining strategy, history, communication, data, ethics, and power in agentic commerce.

Keywords: agentic commerce; antenarrative; true storytelling; qualimetrics; platform governance

Introduction

The arrival of an artificial intelligence agent that can act across websites, accounts, and payment relationships changes more than the speed of shopping. It rearranges who gets to speak, who gets to see, who gets to decide, and who must answer when something goes wrong. The September 2026 dispute around Meta's Muse and Amazon is therefore not merely a product story or a corporate quarrel. It is an emergent field of storying. Amazon's public concern with authorization, security, and conditions of use; Meta's promise of a private personal agent; retailers' differing choices to partner or refuse; and banks' interest in trusted payments are all

attempts to make a future commerce order plausible. Each actor tells a story, but each also performs a material intervention in code, interfaces, terms, identity protocols, and payment rails.

This paper is written as a working paper for the six volumes of the World Scientific Encyclopedia of Business Storytelling. I take seriously the claim that stories do not simply mirror reality. They organize attention, allocate responsibility, and make some actions seem inevitable while other actions become unthinkable. The paper has three aims. First, it develops an integrated theoretical account of agentic commerce as a contested storytelling terrain. Second, it demonstrates a cautious storying and quantification protocol appropriate for a live case where public facts are still incomplete. Third, it offers contributors a practical architecture for articles that move between strategy, history, communication, methods, true storytelling, and postcolonial power without reducing one mountain of storytelling to another.

The study makes a deliberate methodological distinction. A verified public source can establish that Meta introduced Muse as a personal agent, that it is designed to act through a secure virtual machine, and that agentic systems are entering commercial and payment settings. It cannot by itself establish every motive attributed to Amazon, Meta, retailers, or banks. Early reports, corporate announcements, policy statements, and commentary belong to different genres and have different truth claims. The original stimulus for this paper contains analytical propositions that are useful as prompts, including claims about security, advertising, identity, consumer control, and liability. I do not convert those propositions into settled fact. I treat them as antenarrative fragments to be placed beside independently verifiable materials.

This is important because the temptation in AI research is to count first and ask questions later. A large language model can rapidly summarize, classify, and score claims. Yet the result may merely automate the storyteller's own exclusions. The alternative offered here is

qualimetrics: a disciplined movement between qualitative interpretation and transparent measurement. The numbers in a preliminary coding table do not prove the meaning of a conflict. They reveal patterns that require return to the texts, the actors, the missing voices, and the material consequences. In that sense, quantification becomes an occasion for answerability rather than an alibi for it.

The Case as an Unfinished Antenarrative

Meta introduced Muse in September 2026 as a personal AI agent intended to help people pursue goals through a dedicated secure virtual machine and connections to everyday services (Meta, 2026). Contemporary reporting described the system as capable of assisting with tasks such as email, booking, selling, shopping, and payment-related activity, while also reporting continuing concerns about reliability, privacy, and the reach of delegated access (Reuters, 2026). The agent is not a neutral shopping cart. It appears at the crossing of a person's data, a platform's rules, a retailer's commercial interface, and a financial institution's risk obligations.

In such a setting, Amazon's resistance to an unauthorised agent can be read as a security and consumer-protection narrative, but it also operates as an assertion of platform sovereignty. The platform has long been a site where discovery, advertising, fulfillment, and checkout are coupled. An intermediary agent can disaggregate those moments. It can guide attention away from sponsored placement, translate a retailer's interface into a recommendation for the customer, and make the customer relationship more contestable. That does not negate security concerns. It means that security, economic value, and control may be entangled in the same sentence.

Meta's counterstory is one of user delegation and personal assistance. A personal agent promises to make space for a person's goals beyond the interfaces designed by incumbent

platforms. Yet delegation is not emancipation by definition. An agent can only act through permissions, data stores, models, connectors, and commercial arrangements that are mostly invisible to the person it ostensibly serves. Its claim to work for the user must therefore be examined at the level of data practice, consent, reversibility, error repair, and accountability. The conflict has an antenarrative quality precisely because no stable plot has yet won.

Banks enter the story because agentic commerce eventually touches payment authorization, disputed transactions, fraud controls, consumer protection, and the liability chain. Their intervention cannot be treated as purely ethical or purely self-interested. It is both a response to genuine systemic exposure and a narrative attempt to preserve a role in the evolving payment order. A business storytelling approach asks who drafted the principles, which consumers or merchant workers were present, what operational safeguards can be inspected, and how the principles will be revised when they fail.

Theoretical Review

Storytelling Antenarrative and Organization

Organizational storytelling is not an optional communications activity performed after strategy is made. It is part of how organizations coordinate, legitimate, and contest action. Boje's (2001) narrative methods distinguish retrospective narrative from antenarrative, the fragmented and prospective bet on a story that may yet take form. Antenarrative is especially useful for AI cases because an announcement, an API policy, a security warning, a payment rule, and a user complaint may each be a fragment whose relation to the eventual story remains unsettled. A storying analysis follows those fragments before a single corporate plot hardens them into inevitability.

Czarniawska (2004) similarly shows that organizations are constituted through narrative practices, while Bakhtin's (1984) dialogism refuses the fantasy of a final authoritative word. For the Amazon Meta Muse dispute, this means that a corporate statement is not the whole case. It is one utterance in a dialogic field that includes engineers, sellers, consumers, bank compliance teams, regulators, journalists, and systems that answer only through their designed behavior. The analyst's task is to notice both the statement and the response it forecloses.

The term *storying* stresses process. It directs inquiry toward how a narrative is made in real time through documents, interface prompts, authentication flows, marketing images, risk assessments, and conversations. Stories travel through artifacts. A customer-facing notice that an agent is unauthorized is a staged artifact, not just a neutral fact. So is a privacy architecture diagram, a consent screen, an API specification, or a bank's ethics statement. These artifacts authorize particular actors and distribute suspicion, trust, and the burden of proof.

Platform Governance and the Political Economy of Intermediation

Platform scholarship clarifies why the dispute cannot be reduced to a disagreement over a bot. Platforms are infrastructures that coordinate markets while privately governing access, visibility, data, and participation (Gillespie, 2018; Plantin et al., 2018). Their terms of service and technical controls work as rules of a quasi-public space, even though they are issued by private corporations. An agentic intermediary challenges the platform's control over the path between consumer attention and transaction completion.

Srnicek (2017) describes platforms as business models organized around the extraction and control of data. Zuboff (2019) extends the critique by showing how behavioural data can become a source of prediction and intervention. Agentic commerce adds a complex twist. A personal agent may appear to take data power away from a retailer's interface, yet it also

consolidates data and decision power in the agent provider's infrastructure. The question is not whether intermediation disappears. It is which intermediary becomes authorized to narrate consumer preference and to capture its value.

The agent's ability to act makes governance concrete. Identity disclosure, authentication, rate limits, consent scopes, audit logs, revocation, and remediation are not administrative details. They are institutional grammar. They specify who may do what, under which conditions, with what evidence, and with what consequence. The design of these rules is a storytelling act because it embeds a theory of the trustworthy customer, agent, retailer, and bank.

True Storytelling and Moral Answerability

True storytelling offers an ethical test that is broader than factual accuracy. In the framework used here, true storytelling asks whether actors are being true, making room for other stories, creating a plot that gives direction, acting with timing, helping stories along through experimentation, staging coherent artifacts, and reflecting on the value their stories create. The seven principles do not operate as a simple compliance scorecard. They are invitations to examine whether an organization is answerable for the world its story performs.

Bakhtin's (1993) non-alibi in being is central. One cannot escape responsibility by pointing only to the system, policy, model, or market. In AI governance, this means Amazon cannot resolve answerability simply by saying that a system is unauthorized; Meta cannot resolve it by saying that the user authorized a connection; banks cannot resolve it by publishing principles; and users cannot be left holding liability for a system whose operational complexity they cannot see. Answerability is distributed but not dissolved.

Floridi et al. (2018) argue for AI ethics that moves from abstract principles toward practices that are explicable, fair, and sustainable. The storying contribution is to add the

question of whose account of fairness becomes dominant. A principle that looks universal can conceal an incumbent's commercial interest or an innovator's hunger for scale. True storytelling therefore requires evidence, dialogue, and revisability, not merely good language.

Tesseract Leadership Ghost Vortex and Qualimetrics

The tesseract extension brings a four-dimensional question to leadership in AI: task, humanistic concern, antenarrative becoming, and answerability. The framework resists treating leadership as a heroic individual trait. It asks how executive dispositions, corporate scripts, and technical architectures co-produce an organizational field. The fourth dimension is particularly important because an effective task response can still be ethically evasive, and an inclusive public statement can still lack the operational means to repair harm.

Ghost Vortex names the way prior scripts, executive identities, market narratives, and algorithmic layers become palimpsestic in a system. What looks like a fresh AI intervention may carry sedimented assumptions about surveillance, advertising, exclusion, and control. In the case at hand, a security warning may carry the ghost of earlier platform battles over scraping, price comparison, and third-party access. An empowerment claim may carry the ghost of data extraction and platform expansion. The point is not to psychologize individual leaders. It is to trace how institutional character becomes code.

SEAM qualimetrics provides a way to join narrative interpretation with structured inquiry into hidden costs and organizational dysfunction (Savall & Zardet, 2008). A qualimetric analysis can count the frequency of security, autonomy, control, consumer, liability, or collaboration claims, but it must then ask what the count omits. Whose time is spent repairing agent errors? Which workers absorb escalation work? What training and coordination costs are transferred to

merchants, banks, and consumers? Those are hidden costs that a narrow transaction metric will not disclose.

A Six Volume Research Architecture

The encyclopedia can treat agentic commerce as one case that travels across six interlaced volumes. The aim is not six isolated descriptions. It is a research braid: each volume asks a distinct question, while each keeps the other questions visible.

Volume lens	Primary question	Evidence and method	Contribution
Strategy and organization	Who controls the customer path and value capture?	Business models, terms, partnerships, interface analysis	Shows strategic boundary work.
History and institutions	Which earlier struggles are being reactivated?	Genealogy, timelines, archival policy texts	Prevents presentism.
Communication and storying	How are security, autonomy, and trust narrated?	Discourse analysis, antenarrative mapping	Makes silences visible.
Methods and big data	What can be counted without losing context?	Transparent coding, audits, qualimetrics	Connects patterns to hidden costs.
True storytelling and ethics	Who is answerable for consequences?	Seven-principle inquiry, stakeholder dialogue	Tests claims against practice.
Power and plural worlds	Whose knowledge and livelihood are excluded?	Postcolonial, Indigenous, labor, consumer lenses	Reopens the field beyond incumbents.

Table 1

The strategic lens begins with the architecture of value capture. Amazon's resistance, Meta's agent, a retailer's partnership, and a bank's payment rule are strategic choices, but their strategy is not only economic. They determine whether a consumer sees an ad, whether a seller is represented through a platform's taxonomy, whether a third-party agent can compare offers, and whether a payment can proceed. Strategic storytelling in this setting works through legal prose and technical constraint as much as through executive speeches.

The historical lens cautions against the rhetoric of unprecedented disruption. Web crawling, comparison shopping, browser automation, app-store gatekeeping, and payment intermediaries have all produced conflicts over unauthorized access and customer ownership. Historical comparison should not flatten difference. Agentic systems have a new degree of

delegated action and can combine memory, personalization, and transaction authority. But genealogy identifies which corporate arguments are novel and which are old scripts wearing the face of AI.

The communication lens follows the competing genres: announcement, warning, policy, press report, investor discussion, bank principle, and user account. It asks how each genre frames agency. Does it speak of a customer as sovereign, vulnerable, or irresponsible? Does it call an agent a tool, a person-like assistant, a bot, or an unauthorized actor? The words matter because they allocate trust before evidence is examined. The volume invites methods that remain close to wording while also following the material consequences of wording.

The methods volume does not treat data as a neutral corrective to story. Its challenge is to make data collection itself answerable. A coding scheme needs a publicly stated corpus, operational definitions, intercoder discussion, an audit trail, and a clear account of uncertainty. It should never conceal an interpretive judgment under a numerical total. The ethics volume then asks whether the system makes room for those affected, while the power volume asks whether the available categories already carry a Western, corporate, consumerist account of what commerce is for.

Research Design and Transparent Preliminary Quantification

This working paper proposes a preliminary, reproducible protocol rather than claiming a completed large-scale empirical study. The corpus should contain dated public material from clearly identified genres: official product announcements, platform terms and notices, retailer partnership statements, bank or payment-network principles, regulator materials, reputable journalism, and participant accounts. Every item should be assigned a source type, date, actor,

intended audience, and evidentiary status. Verified statements, reported claims, and analytic interpretations must not be merged into one column.

The unit of analysis is the claim-event: a bounded statement linking an actor, a proposition, an object, and a consequence. For example, an official product claim about a secure virtual machine is coded as a Meta architectural claim; it is not coded as proof that all data handling is safe. A platform statement that an agent is unauthorised is coded as a platform access claim; it is not coded as proof of harmful conduct without additional evidence. This small discipline is vital in a fast-moving controversy where repetition can mimic corroboration.

Coding categories include security and privacy, user autonomy, commercial control, identity and authorization, payment liability, consumer protection, collaboration, exclusion, evidence quality, and remedy. A single claim may receive more than one code. Coders should maintain a reflexive memo that records uncertainty, contested interpretation, and missing stakeholder voices. The goal is not to force agreement. It is to make disagreement inspectable.

Coding dimension	Operational question	Illustrative indicator	Interpretive caution
Security and privacy	What risk is named?	Credential, data, fraud, VM, audit	A claim of risk is not proof of harm.
Autonomy	Who may delegate action?	User control, opt-in, revocation	Choice can be constrained by design.
Control and value	Who owns discovery and checkout?	Ads, traffic, fees, access	Economic motive may coexist with safety.
Answerability	Who repairs a failure?	Audit log, liability, escalation	Responsibility can be distributed, not erased.
Making room	Who is absent from the story?	Consumers, workers, small sellers	Absence is an analytic finding.
Reflection	How will claims be tested and revised?	Metrics, pilot, review, remedy	Metrics need stakeholder interpretation.

Table 2

A preliminary count may then calculate the proportion of claim-events per category, per actor, and per source genre. The formula is simple: category salience equals the number of coded claim-events in a category divided by total coded claim-events in the specified corpus. The formula is not a truth score. It is a visibility measure. A high salience of security language could

indicate a genuine risk focus, strategic boundary work, or both. The count tells researchers where to return for close reading and stakeholder inquiry.

Reliability should be treated as a conversation, not a mechanical ritual. Two coders may independently code an initial subset, compare their categories, document reasons for difference, revise the codebook, and retain the original coding memos. Statistical agreement can be reported when a sufficiently stable corpus and coding team exist, but qualitative disagreement must not be discarded. In an antenarrative case, disagreement may reveal the very ambiguity that actors are exploiting.

SEAM adds a second quantification layer: hidden-cost mapping. Researchers can document the time spent by customer-service workers resolving delegated-action errors, the training required for merchant staff, the cost of implementing new authentication, the time consumers spend detecting and contesting errors, and the coordination costs faced by banks. These measures should be accompanied by stories of experience. Otherwise, a hidden cost becomes merely another variable whose human consequence has been abstracted away.

Preliminary Storying Results

The preliminary storying result is a three-way tension rather than a settled binary conflict. The first cluster is protective platform sovereignty: access is framed as conditional, agents are cast as potential risks, and trust is located in the platform's ability to govern its own environment. The second cluster is delegated personal autonomy: the agent is framed as an extension of the person, and value is located in reducing friction across sites and services. The third cluster is accountable payment infrastructure: trust is framed through identity, authorization, control, and liability. These clusters overlap, but they point toward different institutional designs.

The most striking absence is the consumer as a situated person rather than as an abstract beneficiary of safety or convenience. The customer whose address, credentials, shopping history, or disputed charge is involved has a story that cannot be adequately represented by a corporate privacy promise or an access prohibition. A similar absence concerns small sellers and front-line workers. They may experience agentic commerce as new demand, new visibility, new error correction, or new dependency, but their accounts are not normally central to the official storyline.

A second result is the instability of the word authorization. It can mean user permission, platform permission, merchant permission, payment authorization, or legal authority. Those meanings cannot be collapsed. A person may permit an agent to search; a platform may forbid automated access; a bank may require a second confirmation for payment; a merchant may welcome the agent. The political work of the conflict occurs partly by sliding between those meanings. A rigorous paper should define the scope of authorization at every turn.

A third result is the prominence of staging. The secure virtual machine, the warning notice, the consent screen, and the ethics principle are scenographic objects. They make an invisible infrastructure visible in a particular way. They may reassure, deter, or shift blame. A true storytelling analysis asks whether the artifacts are matched by inspectable practice: Is there a meaningful audit trail? Can a person undo an action? Is agent identity verifiable? Is a remedy accessible without specialist knowledge?

Tesseract Reading of the Conflict

A tesseract reading does not rate individual executives from a distance as if leadership were a personality contest. It traces four dimensions in organizational action. Task asks whether actors can make the system work reliably. Humanistic concern asks whether the design respects

affected people and their capacities. Antenarrative becoming asks whether actors can experiment, listen, and change before a conflict hardens. Answerability asks who owns the consequence when the system misleads, excludes, or harms.

Amazon's task narrative emphasizes controlled access and platform rule enforcement. Its humanistic and answerability dimensions require a more demanding question: does a refusal create a safe, transparent pathway for legitimate agentic access, or does it simply protect the incumbent interface? Meta's autonomy narrative scores as an antenarrative opening because it imagines new consumer possibilities. Yet it must be judged on the humanistic and answerability dimensions of disclosure, consent, error repair, data minimization, and independent assessment. Banks may articulate answerability, but their task dimension requires enforceable mechanisms rather than principle alone.

Ghost Vortex deepens this reading. The conflict may be driven not by any single leader's will but by inherited architecture: advertising systems trained to value direct attention, platform policies trained to exclude unapproved intermediaries, and social-media infrastructures trained to convert personal data into engagement. Those ghosts operate through code and organizational habit. The analytic question is how a new agent repeats or interrupts them. A tesseract intervention seeks designs that make the inherited scripts contestable.

Implications for Contributors to the Six Volumes

Contributors can use this framework without writing about Amazon or Meta. The transferable move is to begin with a live organizational controversy and resist its ready-made plot. Collect the competing utterances, identify the material artifacts through which they act, make a timeline of events and silences, and then locate the case across the six volume lenses. A

strong article does not need to cover every lens equally. It should state which mountain it climbs and show how it remains in relation to the others.

For strategy contributors, the question is how narrative frames help firms establish or contest control over value chains, interfaces, data, and customer relationships. For historical contributors, the question is what earlier technical and institutional struggles haunt the present. For communication contributors, the question is how competing words and genres constitute trust. For methods contributors, the challenge is to build transparent, auditable mixtures of close reading, coding, and data analysis. For ethics contributors, the question is answerability in action. For power contributors, the question is whose stories, languages, labor, and worlds are left outside the dominant account.

The framework also invites comparative work. Researchers might compare agentic commerce across retail, travel, health, higher education, or public services. They might compare official narratives with worker, consumer, or community accounts. They might compare a platform that offers a verified agent interface with one that blocks agents entirely. They might trace how a regulatory consultation transforms a voluntary ethics principle into enforceable operational requirements. Such studies can make the encyclopedia a living conversation rather than a catalogue of settled categories.

A methodological invitation is equally important. Authors should publish enough of their codebook, corpus logic, and uncertainty memos for readers to inspect their path. They should state what their figures do not establish. They should avoid treating proprietary model outputs as independent evidence. And they should return findings to people who live with the consequences of the system. This is not a demand for impossible neutrality. It is a demand that the author's own storying be answerable.

Institutional Grammar for Delegated Action

Institutional grammar provides another bridge between storying and system design. Every agentic-commerce arrangement contains an implied statement of the form: an actor may perform an action on an object under specified conditions, using specified evidence, with specified consequences. The actor may be a person, an AI agent, a platform, a merchant, a bank, or a regulator. The action may be search, recommend, authenticate, purchase, disclose, revoke, or repair. The object may be data, an account, a product listing, a payment credential, or a consumer's time. A narrative analysis that does not trace this grammar risks becoming detached from the institutional work that narratives do.

Consider the phrase 'the user authorized the agent.' It is grammatically incomplete. Which user, which agent capability, which duration, which data field, which merchant context, and which remedy? A meaningful authorization is not a blanket checkbox. It has scope, time, legibility, and reversibility. It also has relational consequences. An authorization that is easy for a consumer to grant but difficult to understand or revoke may be formally voluntary while substantively coercive. A storytelling inquiry can make that difference visible by asking how the authorization is staged and what story of the competent user it presumes.

The same is true for platform permission. A platform can lawfully prohibit automated access, yet a business storytelling analysis asks whether its rule makes room for a negotiated, accountable pathway. A blanket exclusion may be a sensible interim response to a genuine threat. It may also protect a business model from an alternative intermediary. The difference lies not solely in motive but in institutional design: public criteria, authenticated identity, evidence standards, pilot arrangements, due process, and review. Such grammar can turn an adversarial story into a testable governance proposition.

Payment grammar is even more exacting. A bank needs to know whether an action is informational, preparatory, or transactional; whether a person confirmed the amount and merchant; whether the agent's identity is cryptographically verifiable; and how a disputed payment can be reconstructed. The presence of an audit log is not enough if a consumer cannot obtain and understand it. The story of consumer control must therefore be instantiated in recognizable action rights: pause, inspect, approve, revoke, contest, and obtain remedy.

Hidden Costs and the Work of Repair

The agentic-commerce debate is often narrated through efficiency. An agent can save the consumer time, route a shopper to a lower price, and lower transaction friction. Those benefits may be real. But SEAM asks researchers to identify the hidden costs that migrate outside the headline transaction. A merchant may need staff to correct inventory mismatches or respond to agents that do not follow ordinary customer flows. A platform may need new detection, identity, and appeals teams. A bank may need investigators, revised risk models, and customer support. A consumer may spend hours disentangling an error made in seconds by a delegated system.

These costs are not merely operational. They have a moral distribution. When a consumer receives a confusing transaction, who bears the burden of proving what happened? When a small seller's listing is misrepresented, who has the capacity to seek correction? When a call-center worker must explain an opaque agent action, what authority and training does that worker have? A system can advertise personalization while externalizing repair to people who possess the least technical and institutional power. This is why qualitative accounts of repair work belong alongside incident counts and cost estimates.

A practical hidden-cost audit can begin with three C's: communication, coordination, and competence. Communication includes warnings, handoffs, explanations, and appeals.

Coordination includes the time required for teams across platform, merchant, agent provider, and bank to resolve incidents. Competence includes training, access to logs, and the authority to act. Researchers can map each C across the lifecycle of an agent action: onboarding, discovery, recommendation, checkout, failure, dispute, and remediation. They can then ask whether the promised time saving for one actor produces uncounted work for another.

The audit should not assume that all hidden costs are negative or that automation is inherently extractive. A verified agent protocol might reduce fraud, lower repetitive work, and make services more accessible. The analytic question is empirical and dialogic: which costs and benefits occur, for whom, under what controls, and with what voice in the redesign process? That question keeps a working paper from becoming either technological celebration or technological refusal.

Plural Worlds and the Limits of the Universal Consumer

The default story of agentic commerce often imagines a universal consumer: a person with stable connectivity, a bank account, a readable consent screen, a predictable address, and the confidence to correct a machine. That consumer is an abstraction. People live in different linguistic, cultural, economic, disability, legal, and relational worlds. An agent that optimizes shopping according to a dominant data profile may marginalize local merchants, nonstandard household arrangements, cash-based practices, or forms of trust that are not legible to platform metrics.

A postcolonial and Indigenous storytelling lens does not add diversity as a decorative final section. It challenges the source of the categories. What counts as a good price, an efficient decision, a legitimate identity, or an authorized transaction? Whose knowledge is used to train recommendations? What happens when a platform's taxonomy translates a community's goods,

stories, or obligations into a standardized product field? These questions are material because agentic systems make selections at scale and can convert local difference into an error condition.

Researchers should therefore seek cases beyond large U.S. platforms and invite methods that are accountable to community protocols. This may include participatory story circles, multilingual corpus work, community review of findings, and refusal to extract stories that are not available for publication. It may include comparing agentic commerce in rural settings, cooperative enterprises, informal markets, or Indigenous-owned businesses. Such work extends the encyclopedia's purpose: business storytelling becomes a meeting ground among mountains, not an imperial map that names every mountain from one summit.

The practical implication is humility. No coding scheme is culturally innocent. Categories such as autonomy, privacy, and consumer benefit must be discussed with the people whose lives they purport to describe. The researcher should document not only what was coded but also what could not responsibly be coded, what was withheld, and what needs a different method. That is an ethics of answerability in research practice.

Future Scenarios and Antenarrative Interventions

Antenarrative analysis is useful because it keeps alternatives alive. One possible future is enclosure: each large platform builds a proprietary agent, blocks rivals, controls identity, and makes customers traverse closed commercial worlds. A second future is delegated federation: independently verified agents can access participating services through common identity, consent, payment, and audit protocols. A third is fractured convenience: agents can act across systems, but error handling, liability, and consumer understanding lag far behind. These are not predictions. They are stories that guide present design choices.

The enclosure scenario may offer clear accountability within one company while constraining competition and user choice. The federation scenario may widen access but requires strong governance, shared standards, and dispute resolution. The fractured-convenience scenario may maximize rapid rollout while shifting hidden costs onto consumers and frontline workers. A storying paper can use these scenarios to ask which actors gain narrative authority in each future, which artifacts make that future likely, and which early interventions could change the path.

An antenarrative intervention is small enough to be tested but consequential enough to alter the plot. Examples include a time-limited sandbox for authenticated agents; a common, machine-readable disclosure standard; a consumer-facing record of every delegated action; an independent incident-review panel; a requirement that material purchases receive meaningful confirmation; and a merchant appeal pathway when an agent misrepresents an offer. None is a final solution. Each creates an experiment through which actors can learn rather than merely trade accusations.

The seven true storytelling principles give criteria for evaluating such experiments. Does the sandbox make room for small merchants and consumer advocates? Does it have a clear plot and an end date? Is its timing appropriate to demonstrated risk? Are its interfaces staged in a way people can understand? Does it include reflection and a published path for revision? An experiment that is inaccessible or opaque merely restages power. An experiment with shared learning can become a different kind of organizational story.

Research Agenda and Propositions

The framework generates propositions for future empirical work. First, the more an actor frames agentic access exclusively through security without publishing inspectable criteria and remedy pathways, the more likely stakeholders will interpret the frame as strategic boundary

work. Second, consumer trust will depend less on broad claims of privacy than on the legibility of delegated actions, revocation, and redress. Third, hidden costs will be concentrated in organizations and roles with weak access to the technical and contractual information needed for repair. These are propositions, not findings; they identify what a multi-site study should examine.

A second set of propositions concerns governance. Shared agent identity and audit standards may reduce conflict only if they are accompanied by meaningful participation from consumers, small merchants, workers, and communities. Voluntary principles without operational tests may contribute to legitimacy but will not reliably establish answerability. Conversely, rigid platform control may reduce certain risks in the short term while discouraging collaborative safety innovation. These propositions invite comparative research across sectors and jurisdictions rather than a single-case verdict.

A rigorous study could combine longitudinal document analysis, interviews, ethnography of repair work, interface walkthroughs, incident records, and participatory workshops. It could publish an evolving corpus and codebook, while protecting material that participants do not wish to disclose. Its quantitative component could include category salience, time-to-remedy, frequency of revocation, appeal outcomes, and the distribution of repair labor. Its qualitative component would return to the stories that make those measures intelligible. This is the promise of qualimetrics: not a compromise between numbers and narratives, but a disciplined relationship between them.

Such a research agenda needs plural authorship. Engineers understand system affordances; legal scholars understand authority and liability; organizational researchers understand coordination; historians understand sedimented conflict; critical theorists understand power; and participants understand the consequences of design. The encyclopedia can create a

place for these voices without dissolving their differences. The task is to organize dialogue that remains open to being changed by what it hears.

Standards for Evidence and Corporate Claims

Agentic-commerce research must keep an evidentiary hierarchy. Official statements establish what an organization says it is doing, not necessarily what the system does in every circumstance. Journalism can document reported events and attributed responses, but its claims should be traced to named evidence where possible. Regulatory documents establish legal requirements or consultation positions, not a system's compliance. User accounts can reveal lived experience and harms, but they require ethical handling and should not be generalized beyond their context. Technical audits, incident records, and reproducible experiments can strengthen causal claims, but they too have scope conditions. The discipline of identifying source type is a safeguard against the rapid circulation of unverified AI narratives.

Corporate claims about privacy, security, and safety require particular care. A claim that an agent uses a secure virtual machine is a claim about an architectural feature; it does not answer questions about connector permissions, data retention, human review, prompt injection, account recovery, or downstream sharing. A claim that an agent does not directly see a password is likewise not a complete account of the risk landscape. Conversely, a platform claim that a third-party agent is unauthorized does not itself establish malicious intent or technical harm. The working paper format is well suited to making these distinctions while the evidence develops.

For this reason, authors should include an evidence-status column in any corpus and label quotations carefully as official claim, reported claim, independent finding, or analytical interpretation. They should avoid circular sourcing, in which several articles repeat one unverified assertion. They should archive web materials because AI product pages and terms

may change without preserving prior versions. And they should treat an inability to inspect proprietary systems as a finding about governance transparency rather than an invitation to fill the gap with speculation.

These standards do not weaken critical inquiry. They make critique more powerful. A precise question—what evidence supports this safety claim, who can inspect it, and who bears the cost if it fails—is harder to dismiss than a generalized accusation. It is also more useful to practitioners who genuinely want to improve a system. The point is not to grant corporations the final word; it is to make every word answerable to the material conditions it names.

Contributor Workflow for an Encyclopedia Article

An author beginning an encyclopedia article can use a seven-step workflow. First, name a live tension rather than a settled topic: for example, delegated diagnosis in health care, AI-assisted hiring, robotized logistics, or automated welfare eligibility. Second, identify the story fragments and artifacts through which the tension is being made. Third, locate the primary volume lens while naming the companion lenses that will remain visible. Fourth, establish a bounded corpus and evidence hierarchy. Fifth, conduct close storying analysis before coding. Sixth, use quantification to test visibility patterns and hidden costs. Seventh, return the interpretation to the ethical questions of answerability, remedy, and plural participation.

The sequence matters. Beginning with a machine-generated codebook may make the case appear tractable but can erase the terms that participants use for themselves. Beginning with close reading gives the researcher a chance to hear unfamiliar metaphors, silences, and contestations. Coding can then be designed to stay accountable to the material. The final return to participants is equally important. An article should identify which claims are provisional, who should be invited to contest them, and what organizational experiments could test the analysis.

Authors should decide early whether their article is conceptual, historical, qualitative empirical, mixed-method, or practice-based. Conceptual work should show the relation among constructs and use a real case as an illustration without overstating empirical claims. Historical work should be rigorous about chronology and archival provenance. Qualitative work should describe sampling, positionality, and analytic procedure. Mixed-method work should show how numbers and narratives inform one another rather than presenting them as unrelated results. Practice-based work should distinguish participant learning from generalizable evidence.

A strong contribution will make its theoretical wager clear. It may argue that an agent is a new narrative actor, that policy is an institutional grammar, that repair work is an overlooked hidden cost, or that true storytelling requires redesign of consent. It should also state what it does not claim. This openness protects the encyclopedia from the appearance of a universal theory while allowing authors to make ambitious, grounded arguments.

Discussion

The case illustrates a broader transformation in organizational communication. When an AI agent acts, it speaks through action as well as text. It can search, rank, select, purchase, disclose, decline, and recommend. Consequently, storying analysis must travel from discourse to affordance. A polished commitment to user control is inadequate if the revocation path is obscure. A security warning is inadequate if the platform offers no inspectable basis for distinguishing harmful automation from legitimate delegated access. The ethical issue is not only what organizations say. It is what their systems make possible and difficult.

The paper also cautions against a simple consumer sovereignty narrative. Delegation can reduce labor and expand access, but it can also deepen dependency on invisible infrastructures. A person may be given an apparently simple choice while being unable to assess the implications

of consent, data retention, model inference, or downstream payments. True storytelling therefore requires a social, not merely individual, account of autonomy. It includes accessible explanation, reversibility, collective governance, redress, and a place for affected communities to change the rules.

There is an additional risk of cognitive surrender. If managers accept an agent's summary, ranking, or score as sufficient, they may lose contact with the lived stories that make the metric meaningful. The same risk faces researchers who use automated text analysis without reading the texts. Qualimetrics is valuable precisely when it interrupts surrender: it makes the measurement process visible, invites interpretation, and sends inquiry back to the situation. The goal is not a perfect score. It is better collective judgment.

Finally, the analysis has limits. The case remains in motion, and some claims reported in the early public conversation may be revised. This paper does not offer findings about private system behavior, internal corporate motives, or the efficacy of untested safeguards. It offers a conceptual and methodological contribution: a way to study a live case without converting news fragments into fact or converting ethics into public relations. The proposed protocol needs application by a diverse research team and engagement with stakeholders who are absent from the current corpus.

Conclusion

Agentic commerce is not simply the next shopping feature. It is a struggle over the story of delegation: who has the right to act for a person, who may govern that action, who receives the value, and who answers for harm. The Amazon Meta Muse bank conflict makes these questions visible because it draws platform sovereignty, personal autonomy, and payment

accountability into the same unsettled field. No party can resolve the conflict by issuing a final monologic story.

The contribution of business storytelling is to keep the field open long enough for responsible institutional design to become possible. Antenarrative identifies the fragments before they harden. True storytelling asks whether room is made for other voices and whether claims are tested in action. The tesseract asks whether task achievement is joined with humanistic concern, becoming, and answerability. Qualimetrics counts without pretending that counting replaces judgment. Together these approaches make it possible to study AI not as a detached technology but as a living, contested organization of relationships.

I invite authors in the World Scientific Encyclopedia of Business Storytelling to use, critique, and extend this architecture. Bring cases from your own mountains. Make explicit the stories that platforms, banks, regulators, workers, communities, and AI systems are performing. Track the hidden costs and the silenced accounts. Offer methods that others can inspect and improve. In the age of AI, the question is never only what the technology can do. It is what story of organization we are allowing it to enact, and whether we will accept responsibility for helping that story become otherwise.

References

Bakhtin, M. M. (1984). *Problems of Dostoevsky's poetics* (C. Emerson, Ed. and Trans.).

University of Minnesota Press. (Original work published 1929)

Bakhtin, M. M. (1993). *Toward a philosophy of the act* (V. Liapunov, Trans.). University of Texas Press.

Boje, D. M. (2001). *Narrative methods for organizational and communication research*. Sage.

Czarniawska, B. (2004). *Narratives in social science research*. Sage.

- European Parliament and Council of the European Union. (2024). Regulation (EU) 2024/1689 laying down harmonised rules on artificial intelligence. Official Journal of the European Union.
- Floridi, L., Cowls, J., Beltrametti, M., Chatila, R., Chazerand, P., Dignum, V., Luetge, C., Madelin, R., Pagallo, U., Rossi, F., Schafer, B., Valcke, P., & Vayena, E. (2018). AI4People—An ethical framework for a good AI society. *Minds and Machines*, 28, 689–707. <https://doi.org/10.1007/s11023-018-9482-5>
- Gillespie, T. (2018). *Custodians of the Internet: Platforms, content moderation, and the hidden decisions that shape social media*. Yale University Press.
- Kitchin, R. (2014). *The data revolution: Big data, open data, data infrastructures and their consequences*. Sage.
- Meta. (2026, September 8). Introducing Muse: The world's first personal AI agent built for everyone. <https://about.fb.com/news/2026/09/introducing-muse-personal-ai-agent/>
- National Institute of Standards and Technology. (2023). Artificial intelligence risk management framework (AI RMF 1.0). U.S. Department of Commerce. <https://doi.org/10.6028/NIST.AI.100-1>
- Organisation for Economic Co-operation and Development. (2019). Recommendation of the Council on artificial intelligence. <https://legalinstruments.oecd.org/en/instruments/OECD-LEGAL-0449>
- Plantin, J.-C., Lagoze, C., Edwards, P. N., & Sandvig, C. (2018). Infrastructure studies meet platform studies in the age of Google and Facebook. *New Media & Society*, 20(1), 293–310. <https://doi.org/10.1177/1461444816661553>

Reuters. (2026, September 8). Meta launches AI agent that can access other apps, send emails and make payments.

Savall, H., & Zardet, V. (2008). Mastering hidden costs and socio-economic performance. Information Age Publishing.

Srnicek, N. (2017). Platform capitalism. Polity.

Zuboff, S. (2019). The age of surveillance capitalism. PublicAffairs.

Appendix A

Minimum Documentation for Replicable Storying Research

At minimum, an empirical article should preserve a dated corpus list; the search or selection rule; a source-genre description; a codebook with examples and exclusions; coding memos; a statement of coder roles and disagreement procedure; a record of category changes; and a clear separation between evidence and interpretation. When automated tools assist with coding, authors should state the model or software, prompt logic, sampling method, human review procedure, and known limitations. Automated output is a research artifact to be checked, not an independent witness.

Where evidence includes people affected by agentic systems, documentation should also record consent procedures, data minimization choices, security protections, and the process for sharing or withholding findings. The researcher should ask whether publication could expose a participant to commercial retaliation, surveillance, or loss of access. These concerns are not a bureaucratic appendix. They are part of the answerability relation that the paper asks other organizations to practice.

Appendix B

Suggested Claim Event Coding Form

Researchers should record each claim-event with: case identifier; date; source URL or archival location; source genre; actor; intended audience; verbatim claim; object of the claim; code or codes; evidence status; implied consequence; absent stakeholder; coder memo; and follow-up source required. Retaining the original wording is essential because paraphrase can prematurely resolve an ambiguity.

Appendix C

Seven Principles Questions for Agentic Commerce

True: What evidence connects the actor's declared values to its operative design and business practice?

Making room: Which affected stories were present when access, identity, and liability rules were written?

Plot: What future does the actor claim to be building, and what trade-offs does it leave unstated?

Timing: Was a safeguard, warning, or launch sequenced with stakeholder readiness and meaningful consultation?

Helping stories along: What reversible pilot, sandbox, or shared standard could be tested before exclusion becomes final?

Staging: Do interfaces, notices, dashboards, and policy documents visibly support the ethical narrative?

Reflecting: What independent feedback, harm reporting, measurement, and remedy process can revise the system?