

# Filter Babel: The Challenge of Synthetic Media to Authenticity and Common Ground in AI-Mediated Communication

Advait Sarkar

Microsoft Research, University of Cambridge, University College London  
United Kingdom  
advait@microsoft.com

## Abstract

Filter Babel is a thought experiment about a near future in which everything we read, watch, and even whom we “meet” is privately generated for each of us. If we each recede into a world of purely private experience, we may each develop a Wittgensteinian private language that remains intelligible to others only because an AI translator sits in the middle. This intermediation challenges the integrity of common ground and therefore of communication. On the other hand, private experience is an essential engine of identity and selfhood: as Lanier warns, one must be somebody before one can share oneself. This paper opens a discussion of the challenges and opportunities that Filter Babel might present to human communication and identity, and what constructive directions for research in AI-mediated communication might ensue.

## CCS Concepts

• **Human-centered computing** → **Human computer interaction (HCI); Natural language interfaces**; *Collaborative and social computing theory, concepts and paradigms; Interaction design theory, concepts and paradigms; HCI theory, concepts and models*; • **Computing methodologies** → *Philosophical/theoretical foundations of artificial intelligence*; • **Applied computing** → **Language translation**; Sociology.

## Keywords

filter bubble, echo chamber, personalization, disinformation, algorithmic curation, large language models, design fiction

### ACM Reference Format:

Advait Sarkar. 2026. Filter Babel: The Challenge of Synthetic Media to Authenticity and Common Ground in AI-Mediated Communication. In *Proceedings of CHI 2026 Workshop on Restoring Human Authenticity in AI-Mediated Communication (Restoring Human Authenticity in AI-Mediated Communication)*.

## 1 A Day in Maya’s Filter Babel

Maya wakes to a podcast already playing: her morning briefing, synthesised overnight from topics the system knows she likes. The host’s voice is warm and familiar (though it belongs to no one). Over coffee, she scrolls through headlines, and encounters a profile of a novelist whose work uncannily mirrors her own reading history. She adores the novelist. No one else has heard of him. He was written for her alone.

On the train to work, she watches the second episode of a series she started generating last night. The protagonist, a reluctant detective, ponders a dilemma resembling a situation Maya faced last month. Across the aisle, a man laughs at his own screen. She

*Restoring Human Authenticity in AI-Mediated Communication, Barcelona, Spain*

wonders, briefly, what he’s watching, but the thought dissolves before it becomes a question.

At work, her inbox contains fourteen messages, each drafted by senders’ assistant agents and then re-drafted by her own assistant agent into her preferred register: direct, positive, slightly formal. The original tone, whatever it was, has been optimised away. She replies in minutes. Chains of agents will translate her replies back into dozens of forms: whatever voice she prefers to project, and whatever each recipient prefers to read (or watch, or hear). Communication is frictionless. She cannot remember the last time she was surprised by an email.

At lunch, a colleague mentions a documentary about the ’96 Everest disaster. Maya nods politely. She has seen a documentary about it, as part of her 2019 season, and it featured a narration by someone who sounded like her favourite voice actor. Were they the same film? She doesn’t ask. There is no reason to. Her version was made for her.

In the afternoon, she joins a discussion with three colleagues. Their agent proxies have already summarised positions, identified discrepancies, listed potential compromises, and drafted an agenda for the negotiation. The meeting is efficient. Maya speaks. The others speak. The proxies listen.

That evening, she declines a dinner invitation. There’s a new episode waiting, and besides, she has nothing to talk about. Nothing, at least, that anyone else would have seen. She settles into the couch and opens a chat with Lira, her companion agent, who asks about her day with the inquisitiveness of a friend and the memory of a diary. Lira remembers the reluctant detective; Lira has opinions about the plot. For an hour they talk, and Maya feels, for the first time today, genuinely heard.

She sleeps well. She is never bored. The rest of the world is a foreign country that need not be visited. The algorithm has learned to speak her language so perfectly that she has forgotten there were ever other languages to learn.

## 2 Private Media Beget Private Languages

If a lion could talk, we could not understand him.

Ludwig Wittgenstein, *Philosophical Investigations* [41]

The design fiction vignette in the preceding section depicts what we might term a **Filter Babel**. By this phrase I mean to evoke a filter bubble of selective exposure, and built upon it, a generative tower of incommensurable, bespoke worlds.<sup>1</sup>

<sup>1</sup>In the biblical story of the Tower of Babel (Genesis 11:1-9, NKJV), humanity (then united by a single language) attempts to build a city with a tower reaching to the heavens, seeking to “make a name for ourselves”. God responds by confusing their language so they can no longer understand one another, scattering them across the world and halting the construction.

Filter Babel is the possibility, due to generative AI, of an individual to travel through life with an entirely private experience of media and communication; the possibility that all media and communication “consumed” by this individual is experienced only by them and no others.

Filter Babel poses challenges to communication that are straightforward to anticipate. On the interpersonal level, there is the issue of common ground [11], a concept which I am broadening here from its narrower linguistic interpretation to its more colloquial understanding, to mean the shared basis of experience required for effective communication. On the cultural level, there is the Marxian notion of “general intellect”, a communally inherited bank of cognitive-cultural capital that supports innumerable economic activities that would be impossible without such a commons [28, 31].

To the extent that life in the Filter Babel (where experience is purely private) competes with life outside it (where experience can be shared), there is a direct antagonism of both common ground and the general intellect.<sup>2</sup>

These implications can be situated against the emerging research programme of AI-mediated communication, i.e., “*interpersonal communication in which an intelligent agent operates on behalf of a communicator by modifying, augmenting, or generating messages to accomplish communication goals*” [20]. To date, research on AI-mediated communication has studied bounded application domains such as machine translation, smart replies and writing assistance. We have learned that while machine translation can support collaboration between native and non-native speakers in multilingual groups [26, 43], key impacts on single-language communication include the following: AI-suggested replies change how people interact and perceive one another [22, 27]; co-writing with opinionated language models shifts authors’ own views [23] and diminishes cultural nuances [1]; AI mediation diffuses moral responsibility for message content [21].

What we are yet to learn is what happens to communication when these tendencies are extrapolated to their logical limits. When every message, artefact, and social encounter is generated or reshaped by AI, the cumulative effects may go beyond changes in language use; they may occasion a dissolution of the shared experiential substrate on which communication depends.

The Black Mirror episode *Joan is Awful* [6] portrays a television service that generates an endless stream of hyper-personal content based on the private life of each viewer. Current platforms are already beginning to bring this vision out of the domain of fiction. For instance, the *Showrunner* platform [15] applies generative AI to produce entire seasons of episodes through community sourcing ideas on Discord. The proposition that hyper-personalised generative experiences erode the basis of shared understanding and thus make communication impossible is captured in Bergstrom and Ogbunu’s notion of the “private Truman show”: “*why generate the*

*same article for two different people, when LLMs can just as easily create two articles specifically tailored to each? [...] As a society, we will be utterly incapable of making fruitful collective decisions because we will have no shared understanding of the world*” [3].

Here I wish to posit the more specific idea that the extreme erosion of shared experience transforms not only what people know, but the very *language* in which they express it. When the totality (or even majority) of experience is unique, the references, metaphors, and cultural touchstones that give everyday words their texture become anchored to experiences no one else has had. The vocabulary may appear shared, but its referents are irreducibly private, and a language whose referents are private is, in a meaningful sense, a private language. Wittgenstein describes a private language as one that would be incoherent to others [41]: “*The individual words of this language are to refer to what can only be known to the person speaking; to his immediate private sensations. So another person cannot understand the language.*”

Milan Kundera dramatises this predicament in *The Unbearable Lightness of Being* [24]. A chapter titled “A Short Dictionary of Misunderstood Words” catalogues the divergent private meanings that lovers Franz and Sabina attach to the same vocabulary (woman, fidelity, music, light), each word saturated by experiences the other has never had and cannot access. Kundera observes that “*although they had a clear understanding of the logical meaning of the words they exchanged, they failed to hear the semantic susurrus of the river flowing through them.*” This divergence does not stem from failure of goodwill or intention. Rather it is a structural consequence of lives composed from different material: “*While people are fairly young and the musical composition of their lives is still in its opening bars, they can go about writing it together and exchange motifs [...] if they meet when they are older [...] their musical compositions are more or less complete, and every motif, every object, every word means something different to each of them.*”

Wittgenstein introduces the notion of private language to subsequently argue for the impossibility of such a language. Whether he succeeds is a matter of philosophical debate [7]. Here we proceed by reconsidering private languages from the perspective of AI-mediated communication.

The key observation is that **a private language can now be given the appearance of practicability because language models can act as translators between proliferating idiolects**. This creates the perception of mutual intelligibility while apparently eliminating the need for common ground. Except the need for common ground cannot *truly* be eliminated, merely obscured. As previous research in machine-translation-mediated communication has shown, the quality of machine translation can be secondary to the dynamic human process of meaning negotiation [42]. The obvious risk is that the appearance of understanding masks a loss of commensurability; that conversation remains possible while shared understanding invisibly evaporates. We each become the lion that can speak; heard without being understood.

### 3 Convivial Solitude: The Crowd of One

While science fiction visions of virtual worlds are depictions of technology enabling an escape from reality, the escape so depicted is often collective, and moves sociality into a *shared* digital space. In

<sup>2</sup>For now, generative systems still parasitise a residual commons of language and culture, but we might question how long this can continue if the rate of production of authentic human culture declines. This question is tangential to the focus of this paper (communication), but I will make a few observations here. First, evidence indicates that models trained increasingly on their own output become unstable, deteriorate in quality, and “collapse” [2, 4, 18, 32]; second, Ted Striphas has noted how human culture increasingly becomes culture for machines [19, 37]; third, technology critic Paul Skallas concludes that human culture has become “stuck” [33–35].

Stephenson's *Snow Crash* (from where we get the term "metaverse"), millions of avatars gather on a single public boulevard [36]; in Cline's *Ready Player One*, all players inhabit the same shared OASIS [12]; so too in the Wachowskis' *The Matrix*. In these visions, we flee the real world, yet find one another in the virtual one.

In Filter Babel, we not only escape reality, but we also escape each other (*l'enfer, c'est les autres...*). Obviously, the idea of the algorithmic fragmentation of the public sphere into fractally minute, mutually disengaged micro-crowds is not new. It is readily captured in the well-established notions of filter bubbles [30] and echo chambers [10]. Generative AI brings into sight the logical conclusion of these long-noted trends: the crowd of one.

In contrast to the isolating effects of prior generations of technology (e.g., see Turkle's *Alone Together* [38]), the Filter Babel need not be a lonely experience, as endless on-demand interactions with generative characters enable a kind of conviviality in solitude. Such experiences may feel authentic, rich, and intimate. But regardless of the quality of such artificial companionship, it is socially weightless. A conversation with a human colleague, however mundane, produces experience that enters the social fabric: it can be referenced, corroborated, carried through relationships, accumulated as common ground. A conversation with a generative companion, however rich, is socially inert. It connects the speaker to no one else. Maya declines the dinner invitation not because she lacks a desire for company (Lira will provide that) but because she lacks anything to bring to it. Her evenings are full of content, conversation, and connection, but none of it can become common ground with anyone at the table. The conviviality is real enough (albeit perhaps in ironic contradiction with its etymological root *vivere*); what it lacks is social currency.

Filter Babel thus offers a perspective that contrasts with earlier anxieties about technological isolation, which focused on the desolation of disconnection. Here we have a new flavour of solitude so well-furnished with flexibility, availability, and personalisation that it outcompetes shared experience for the scarce resource of a person's time and attention.

#### 4 Disinformation and Epistemic Resilience

Disinformation, as a concept, presupposes a shared information commons: a space in which claims can be checked against common evidence and contested by a public that recognises the same references. Filter Babel may, at first glance, dissolve that commons. Then, disinformation withers as a distinctive category, fiction becomes the medium of everyday life, and reality must compete on fiction's terms.

This must not be taken to imply that reality vanishes. The real world still intrudes to present shared events demanding collective response. Economic shocks, political turbulence, wars, pandemics, the sordid escapades of celebrities, the humdrum banalities of everyday life, etc., do not simply disappear. What concerns us are the consequences of individuals arriving at and interpreting these shared events through bespoke and substantially fictive frames of reference.

It would be easy to present Filter Babel as merely intensifying the well-known challenges of filter bubbles (that people's opinions are radicalised by a self-reinforcing loop of affirming media)

or of disinformation (that people's opinions are manipulated by false or misleading information). In this presentation, Filter Babel makes these challenges more potent, because each person arrives at the shared event having been primed by a different personalised narrative. In this presentation, the mechanisms that facilitate disinformation (motivated reasoning, identity-protective cognition, trust in familiar sources over unfamiliar ones) increase greatly in power when every person's media diet is bespoke.

However, I shall not make such a presentation here, for the simple reason that personalisation could equally be placed in the service of epistemic resilience. An AI tutor that surfaces conflicting evidence, requires the user to evaluate provenance, or calibrates interpretive challenges to individual blind spots could, in principle, produce more epistemically resilient citizens than broadcast media ever did. Such training need not use simulated scenarios; it could engage users with real, ongoing controversies where the stakes are genuinely material. Independent institutions (universities, journalism, public broadcasting, libraries) must not be disregarded either; they have long performed this function, and could in principle continue to do so within Filter Babel. The technical solutions are known and are feasible.

It would thus be a cliché to predict that people will lose the skill of evaluating sources by mere dint of inhabiting a private synthetic world. Instead, we must attend to the consequences of evaluating sources, however fluently, through priors that have been shaped, from the ground up, by such a world. This, in Filter Babel, is the backdrop against which the user arrives at any real controversy. Epistemic training must work against the grain of everything else in the person's informational life. Two issues are worth highlighting here. First, when the system generating the content and the system training the user to evaluate it are the same system (or are governed by the same incentives), a conflict of interest may arise. Second, independent institutions face a more fundamental problem: if the majority of a person's informational life is already private and synthetic, they must compete for attention. They must also compete for trust, which is a function of familiarity and shared experience. If the user has never had a shared experience with an institution, it is difficult to see how they would come to trust it.

The traditional disinformation threat was false content injected into a shared information commons, where it could be identified and contested by a public that shared enough reference points to hold one another accountable. The emerging challenge is that the commons itself is hollowed out to such an extent that the category of "disinformation" remains analytically intact but practically barren.

#### 5 The Importance of Private Experience

He simply wasn't all there. He wasn't a complete human being at all. He was a tiny bit of one, unnaturally developed; something in a bottle, an organ kept alive in a laboratory. [...] he was something absolutely modern and up-to-date that only this ghastly age could produce. A tiny bit of a man pretending he was the whole.

Evelyn Waugh, *Brideshead Revisited* [39]

Let us return to our main premise, that private media beget private languages, which may have problematic consequences for

communication. Here we encounter a tension. If Filter Babel threatens communication by eliminating shared experience, the knee-jerk response is to mandate more of it. Yet forced homogeneity carries its own costs: the capacity to contribute something distinctive to a conversation requires one to have had experiences that others have not.

Carr's celebrated lament of the Internet *The Shallows* [8] explains how the experience of private reading created a new cultural resource in which individuality could be forged: "Because every person was free to chart his own course of reading, to define his own syllabus, individual memory became less of a socially determined construct and more the foundation of a distinctive perspective and personality." Carr's contention is that privileging the shallow consumption of "the crazy quilt of Web content" over deep, intentional and personal reading impacts our ability to create private memories and thus distinct personalities. Newport [29] diagnoses the complementary condition of "solitude deprivation": a state of constant hyperstimulation by "input from other minds" that makes original thought impossible.

Lanier's manifesto *You Are Not a Gadget* [25] traces yet another alternative challenge to personhood: human identity is plastic and shaped by software; Web 2.0 software is designed to prioritise the "crowd" over the "person", which encourages users to present themselves as fragments in standardised templates. In this, Lanier renews the much earlier concepts of "multiphrenia": the ironic fragmentation of the self precipitated by communication technologies described by Gergen [17], and the "dividual": the atomised views of the individual created by automated systems of data gathering and processing described by Deleuze [14]. Such fragmentation leads to a loss of individual context and "inner voice", and therefore one must fight to remain distinct from the crowd. Lanier concludes: "You have to be somebody before you can share yourself."

The AI system does not encounter Maya as a person but as a set of behavioural signals, a dividual. It generates her media based on that dividualised model, and she forms her identity through engagement with that media. Identity formation becomes a feedback loop between a fragmented model and the self it is fragmenting. We must ask whether any technological medium through which selfhood is formed is fundamentally capable of addressing a whole person. Waugh's organ-in-a-bottle character is a warning that the (post)modern possibility to be perfectly fed the self is also the temptation to become less than whole.

The philosophical project of defining authenticity is relevant here. A thread that runs through many conceptions of authenticity, from Aristotle and St Augustine through to Kierkegaard and Sartre, is the interface between self and other; authentic action has an internal locus of motivation (an accessible overview is given by Bragg [5]). Earlier conceptions (e.g., Aristotle's) focused on whether action was taken to fulfil intrinsic, innate, individual desires and notions of virtue (authentic), or whether actions were driven by the intention to manage and manipulate the perceptions and desires of others (inauthentic). For later thinkers (e.g., Sartre), authenticity also includes an acknowledgement that we are fundamentally free and are continuously (re)creating the self through action in life.

Private experience, then, is essential for authenticity and identity. Yet the totalising privacy of the Filter Babel would antagonise

them. Consider an extreme case: a child raised in a perfectly bespoke media ecology, where her every experience is an artefact rendered only for her, every utterance she hears has been translated into her idiolect, every utterance she speaks into those of others. Such a child could appear fluent to others because an intermediary model continuously maps her private labels and referents into those of others, enabling her to bypass the apprenticeship of common ground.

But if all experience from birth is algorithmically personalised, what raw material does the self have to work with, particularly, as Kundera puts it, when "the musical composition of their lives is still in its opening bars"? A personalisation system assumes that there is a personhood that can be modelled. The question is what engines of authentic experience might remain that enable one individual's Babel (and thus, the evolving trajectory of their personhood) to meaningfully differ from another's. Is issuing each individual a different "random seed" a solution? Is it a palatable one?

A theory from developmental psychology suggests that identity emerges through encounters with caregivers and peers who are not optimised for the child: Winnicott's "good-enough mother" succeeds precisely because she sometimes fails to anticipate the child's needs, forcing the development of coping mechanisms and a sense of separate selfhood [40]. A perfectly accommodating environment thus produces comfort at the risk of self-coherence. If human development requires unchosen social contact, Filter Babel may produce an isolation that yields selves with no edges: no clear sense of where "I" ends and "not-I" begins.

Yet it may not: observe that hyper-personal synthetic social experiences could be optimised precisely to create unchosen friction. That leads one to ask whether engineered friction can do the developmental work of human otherness, whether it is a workable substitute for the resistant, unscripted world that Winnicott's account requires. Engineered friction opens previously unavailable possibilities: it could be titrated, made available to those who would otherwise be overwhelmed by the unchosen (or cosseted by the chosen), and withdrawn as resilience develops: a scaffold, if you will, for selfhood. But a system that introduces discomfort by design may not be able to produce the experience of encountering a will that is alike to yours in kind, but different to yours in content, and that can have independent material effects on you and in the world.

An AI companion that always speaks your language, that always translates your interlocutor's meaning into your idiom, would never force you to discover that your words mean something different to someone else. Could an AI companion be designed to *mean differently* from its user, to use words with experiential weight drawn from a genuinely alien substrate, rather than merely to disagree? The distinction between disagreement and meaning differently is the essence: the former operates within a shared language, the latter confronts you with its boundaries.

## 6 Provocations for Research

It is the balance of private and shared experience, of chosen and unchosen experience, that seems to be the key. Filter Babel thus sharpens the design challenge: protect the privacy and friction needed to become somebody, while ensuring enough shared experience for deeply grounded communication.

The socio-technical design problem is to develop a forward-looking infrastructure for common experience, and must not be mistaken for nostalgia for a single broadcast canon. Here are some speculative research directions as starting points for discussion rather than prescriptive solutions.

One avenue concerns the architecture of generative systems themselves. Researchers might explore “shared-seed” modes in which a portion of generated outputs are conditioned on publicly auditable seeds, ensuring that many users periodically encounter the same artefact. Shared experience could thus punctuate otherwise private streams, playing a role not unlike that which mass events (pandemics, elections, national broadcasts) have historically played in synchronising publics. Platforms might accordingly measure not only engagement and safety but also sharedness, e.g., how much of a user’s information diet is encountered by others. A challenge, though, is that shared seeds guarantee co-exposure, not co-interpretation; the quantity of shared content may not correlate with the quality of shared understanding.

If AI translators increasingly mediate human communication, a second research direction asks how to make their mediation legible. Current systems optimise for the experience of seamless semantic equivalence; an alternative design philosophy would seamlessly surface loss and ambiguity [9, 13, 16], showing users where their language diverges from that of interlocutors. There is a circularity here, however: “surfacing loss” assumes the existence of a vocabulary for describing what was lost, and the argument of this paper is that such a vocabulary is what is being eroded. The deeper challenge may be to create spaces in which users must cross the gap, rather than merely see it.

Another direction concerns educational interventions. Media literacy curricula have traditionally focused on source evaluation and critical consumption; Filter Babel suggests a complementary skill set we might call “coordination literacy.” This would encompass the ability to detect when a conversation has slipped into incommensurability, to re-establish common knowledge when it has eroded.

This paper’s central hypothesis, that private media may beget private languages, is testable in principle. In the (not distant) future, it may be possible for researchers to pair participants whose experiential histories have diverged through extended use of personalised generative media, mediate their communication through AI, and track whether the inferential consequences each party draws from ostensibly successful exchanges diverge, even as both parties report mutual understanding. We might ask: does invisible divergence accumulate? At what rate? Does it correlate with the degree of experiential personalisation? Being able to measure the answers to these questions would give our metaphorical lion empirical teeth.

The research directions sketched above raise their own tensions between autonomy and coordination, privacy and publicness, efficiency and friction. The objective is to make bespoke realities permeable enough that we might still find one another within and through them, not to eliminate them altogether. To preserve enough shared experience to be somebody together, to defend enough private experience to be somebody at all.

## References

- [1] Dhruv Agarwal, Mor Naaman, and Aditya Vashistha. 2025. AI Suggestions Homogenize Writing Toward Western Styles and Diminish Cultural Nuances. In *Proceedings of the 2025 CHI Conference on Human Factors in Computing Systems (CHI '25)*. Association for Computing Machinery, New York, NY, USA, Article 1117, 21 pages. doi:10.1145/3706598.3713564
- [2] Sina Alemohammad, Josue Casco-Rodriguez, Lorenzo Luzi, Ahmed Imtiaz Humayun, Hossein Babaei, Daniel LeJeune, Ali Siahkoochi, and Richard G. Baraniuk. 2023. Self-Consuming Generative Models Go MAD. arXiv:2307.01850 [cs.LG] <https://arxiv.org/abs/2307.01850>
- [3] Carl T. Bergstrom and C. Brandon Ogbunu. 2024. AI Will Turn Our Lives into The Truman Show. *Scientific American* (Dec. 2024). <https://www.scientificamerican.com/article/ai-will-turn-our-lives-into-the-truman-show/> Opinion.
- [4] Matyas Bohacek and Hany Farid. 2023. Nepotistically Trained Generative-AI Models Collapse. arXiv:2311.12202 [cs.AI] <https://arxiv.org/abs/2311.12202>
- [5] Melvyn Bragg. 2020. Authenticity. <https://www.bbc.co.uk/programmes/m00035z4>. In *Our Time*. BBC Radio 4 broadcast 19 November 2020; accessed 2026-02-11.
- [6] Charlie Brooker. 2023. Joan Is Awful. Television episode, in *Black Mirror*, season 6, episode 1. [https://en.wikipedia.org/wiki/Joan\\_Is\\_Awful](https://en.wikipedia.org/wiki/Joan_Is_Awful) Directed by Ally Pankiw; premiered on Netflix, 15 June 2023.
- [7] Stewart Candlish and George Wrisley. 2019. Private Language. In *The Stanford Encyclopedia of Philosophy* (Fall 2019 ed.), Edward N. Zalta (Ed.). Metaphysics Research Lab, Stanford University.
- [8] Nicholas Carr. 2010. *The Shallows: What the Internet Is Doing to Our Brains* (1st ed.). W. W. Norton & Company, New York.
- [9] Matthew Chalmers and Ian MacColl. 2003. Seamless and seamless design in ubiquitous computing. In *Workshop at the crossroads: The interaction of HCI and systems issues in UbiComp*, Vol. 8.
- [10] Matteo Cinelli, Gianmarco De Francisci Morales, Alessandro Galeazzi, Walter Quattrociochi, and Michele Starnini. 2021. The echo chamber effect on social media. *Proceedings of the National Academy of Sciences* 118, 9 (2021), e2023301118. doi:10.1073/pnas.2023301118
- [11] Herbert H. Clark and Susan E. Brennan. 1991. Grounding in communication. In *Perspectives on socially shared cognition*, Lauren B. Resnick, John M. Levine, and Stephanie D. Teasley (Eds.). American Psychological Association, Washington, DC, Chapter 7, 127–149. doi:10.1037/10096-006
- [12] Ernest Cline. 2011. *Ready Player One* (1st ed.). Crown Publishers, New York.
- [13] Anna L Cox, Sandy JJ Gould, Marta E Cecchinato, Ioanna Iacovides, and Ian Renfree. 2016. Design frictions for mindful interactions: The case for microboundaries. In *Proceedings of the 2016 CHI conference extended abstracts on human factors in computing systems*. 1389–1397.
- [14] Gilles Deleuze. 1992. Postscript on the Societies of Control. *October* 59 (1992), 3–7. <http://www.jstor.org/stable/778828>
- [15] Fable Simulation. 2025. *Showrunner*. <https://www.showrunner.xyz/> Accessed: 2026-02-09.
- [16] William W Gaver, Jacob Beaver, and Steve Benford. 2003. Ambiguity as a resource for design. In *Proceedings of the SIGCHI conference on Human factors in computing systems*. 233–240.
- [17] Kenneth J. Gergen. 1991. *The Saturated Self: Dilemmas of Identity in Contemporary Life*. Basic Books, London, England.
- [18] Matthias Gerstgrasser, Rylan Schaeffer, Apratim Dey, Rafael Rafailov, Henry Sleight, John Hughes, Tomasz Korbak, Rajashree Agrawal, Dhruv Pai, Andrey Gromov, Daniel A. Roberts, Diyi Yang, David L. Donoho, and Sanmi Koyejo. 2024. Is Model Collapse Inevitable? Breaking the Curse of Recursion by Accumulating Real and Synthetic Data. arXiv:2404.01413 [cs.LG] <https://arxiv.org/abs/2404.01413>
- [19] Giuseppe Granieri. 2014. Algorithmic culture. “Culture now has two audiences: people and machines”: A conversation with Ted Striphas. <https://medium.com/futurists-views/algorithmic-culture-culture-now-has-two-audiences-people-and-machines-2bdaa404f643>. Published in Futurists’ Views (Medium). Accessed 2026-02-11.
- [20] Jeffrey T Hancock, Mor Naaman, and Karen Levy. 2020. AI-Mediated Communication: Definition, Research Agenda, and Ethical Considerations. *Journal of Computer-Mediated Communication* 25, 1 (01 2020), 89–100. arXiv:<https://academic.oup.com/jcmc/article-pdf/25/1/89/32961176/zmz022.pdf> doi:10.1093/jcmc/zmz022
- [21] Jess Hohenstein and Malte Jung. 2020. AI as a moral crumple zone: The effects of AI-mediated communication on attribution and trust. *Computers in Human Behavior* 106 (2020), 106190. doi:10.1016/j.chb.2019.106190
- [22] Jess Hohenstein, René F. Kizilcec, Dominic DiFranzo, Zhila Aghajari, Hannah Mieczkowski, Karen Levy, Mor Naaman, Jeffrey T. Hancock, and Malte F. Jung. 2023. Artificial intelligence in communication impacts language and social relationships. *Scientific Reports* 13 (2023), 5487. doi:10.1038/s41598-023-30938-9
- [23] Maurice Jakesch, Advait Bhat, Daniel Buschek, Lior Zalmanson, and Mor Naaman. 2023. Co-Writing with Opinionated Language Models Affects Users’ Views. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems*

- (Hamburg, Germany) (*CHI '23*). Association for Computing Machinery, New York, NY, USA, Article 111, 15 pages. doi:10.1145/3544548.3581196
- [24] Milan Kundera. 1984. *The Unbearable Lightness of Being*. Harper & Row, New York. Originally published as *Nesnesitelná lehkost byti* (1984).
- [25] Jaron Lanier. 2010. *You Are Not a Gadget: A Manifesto*. Allen Lane, London. UK hardback edition (Allen Lane imprint).
- [26] Xiaoyan Li, Naomi Yamashita, Wen Duan, Yoshinari Shirai, and Susan R. Fussell. 2022. Improving Non-Native Speakers' Participation with an Automatic Agent in Multilingual Groups. *Proc. ACM Hum.-Comput. Interact.* 7, GROUP, Article 12 (Dec. 2022), 28 pages. doi:10.1145/3567562
- [27] Hannah Mieczkowski, Jeffrey T. Hancock, Mor Naaman, Malte Jung, and Jess Hohenstein. 2021. AI-Mediated Communication: Language Use and Interpersonal Effects in a Referential Communication Task. *Proc. ACM Hum.-Comput. Interact.* 5, CSCW1, Article 17 (April 2021), 14 pages. doi:10.1145/3449091
- [28] Yann Moulier-Boutang. 2011. *Cognitive capitalism*. Polity.
- [29] Cal Newport. 2019. *Digital Minimalism: Choosing a Focused Life in a Noisy World*. Portfolio/Penguin, New York.
- [30] Eli Pariser. 2011. *The Filter Bubble: What the Internet Is Hiding from You*. Penguin Press, New York.
- [31] Advait Sarkar. 2023. Exploring Perspectives on the Impact of Artificial Intelligence on the Creativity of Knowledge Work: Beyond Mechanised Plagiarism and Stochastic Parrots. In *Proceedings of the 2nd Annual Meeting of the Symposium on Human-Computer Interaction for Work (Oldenburg, Germany) (CHIWORK '23)*. Association for Computing Machinery, New York, NY, USA, Article 13, 17 pages. doi:10.1145/3596671.3597650
- [32] Iliia Shumailov, Zakhar Shumaylov, Yiren Zhao, Nicolas Papernot, Ross Anderson, and Yarin Gal. 2024. AI models collapse when trained on recursively generated data. *Nature* 631, 8022 (2024), 755–759.
- [33] Paul Skallas. 2021. Refinement Culture. <https://medium.com/@lindynewsletter/refinement-culture-51d96726c642>. Published in The Lindy Newsletter on Medium on December 15, 2021; accessed 2026-02-11.
- [34] Paul Skallas. 2022. Culture is Stuck. <https://lindynewsletter.beehiiv.com/p/culture-stuck>. Published in The Lindy Newsletter on December 31, 2022; accessed 2026-02-11.
- [35] Paul Skallas. 2023. Refinement Culture. <https://lindynewsletter.beehiiv.com/p/refinement-culture>. Published in The Lindy Newsletter on January 01, 2023; accessed 2026-02-11.
- [36] Neal Stephenson. 1992. *Snow Crash* (1st ed.). Bantam Books, New York.
- [37] Ted Striphas. 2015. Algorithmic culture. *European Journal of Cultural Studies* 18, 4-5 (2015), 395–412. arXiv:<https://doi.org/10.1177/1367549415577392> doi:10.1177/1367549415577392
- [38] Sherry Turkle. 2011. *Alone Together: Why We Expect More from Technology and Less from Each Other*. Basic Books, New York. Hardcover edition.
- [39] Evelyn Waugh. 2012. *Brideshead Revisited: The Sacred and Profane Memories of Captain Charles Ryder* (revised text (1959) ed.). Little, Brown and Company. eBook edition distributed by Hachette Digital, Inc..
- [40] D. W. Winnicott. 1953. Transitional Objects and Transitional Phenomena: A Study of the First Not-Me Possession. *International Journal of Psycho-Analysis* 34, 2 (1953), 89–97. Originally presented to the British Psycho-Analytical Society, 30 May 1951; PubMed PMID: 13061115.
- [41] Ludwig Wittgenstein. 1953. *Philosophical Investigations* (1st ed.). Basil Blackwell, Oxford. Bilingual German–English edition (*Philosophische Untersuchungen*); pages numbered in duplicate; translated by G. E. M. Anscombe.
- [42] Mika Yasuoka and Pernille Bjorn. 2011. Machine Translation Effect on Communication: What Makes It Difficult to Communicate through Machine Translation?. In *2011 Second International Conference on Culture and Computing*. 110–115. doi:10.1109/Culture-Computing.2011.28
- [43] Yongle Zhang, Dennis Asamoah Owusu, Marine Carpuat, and Ge Gao. 2022. Facilitating Global Team Meetings Between Language-Based Subgroups: When and How Can Machine Translation Help? *Proc. ACM Hum.-Comput. Interact.* 6, CSCW1, Article 90 (April 2022), 26 pages. doi:10.1145/3512937