Information Technology & People

When programmes disagree about programmes: large language models, code poetry and the problem of programme–programme communication

Received 3 February 2026
Revised 26 March 2026
1 June 2026
15 July 2026
21 July 2026
Accepted 21 July 2026

Steffen Roth
CERIIM, Excelia Business School, La Rochelle, France;
Wolfson College, University of Cambridge, Cambridge, UK and
Centre for Research in the Arts, Social Sciences and Humanities, University of Cambridge, Cambridge, UK, and
Vincent Lien
Clare College, University of Cambridge, Cambridge, UK and
Faculty of Education, University of Cambridge, Cambridge, UK

Abstract

Purpose – This article examines the theoretical implications of organisational decisions being increasingly prepared, filtered, produced through or attributed to computational programmes rather than individual human actors. It asks under what conditions interactions between such programmes may be analysed as communicatively consequential and what this implies for management and organisation theory.

Design/methodology/approach – The article develops a conceptual analysis grounded in Luhmannian social systems theory. It uses a deliberately simplified, user-moderated exchange between two large language models, ChatGPT and Gemini, as an illustrative vignette. Their interpretations of code poems serve as a heuristic device for examining interpretive divergence, role formation and double contingency.

Findings – The illustrative exchange suggests that mutual responsiveness between computational programmes need not produce interpretive convergence. Instead, recursive reference to one another's outputs may stabilise distinct observational positions and patterned forms of mutual irritation. The article does not treat the vignette as an empirical test of programme–programme communication. It uses it to develop the conceptual proposition that large language models may be observed as organisationally instantiated decision programmes whose outputs become communicatively consequential when they are recursively incorporated into further organisational decisions.

Originality/value – The article contributes to management and organisation theory by reframing artificial intelligence not as a tool or agent but as a decision programme operating within organisational autopoiesis. By introducing programme–programme communication as an analytical lens, it shows how generative information technologies may create new decision premises, stabilise differentiated roles and reshape organisational coordination, responsibility and control in programme-rich environments.

Keywords Programme-programme communication, Organisational autopoiesis, Double contingency, Artificial intelligence, Code poems

Paper type Conceptual paper

1. Introduction: communicating decision programmes

It has become a commonplace observation that organisational decision-making is increasingly computer-mediated and automated (Raisch and Krakowski, 2021). From algorithmic scoring systems and rule-based compliance checks to predictive analytics and generative artificial intelligence, organisational decisions are ever more frequently prepared, filtered, or even executed through computational procedures (Agrawal *et al.*, 2022; Cingillioglu and Schoettner, 2025; Curchod *et al.*, 2020; Kelan, 2024). In management and organisation studies, this development is often discussed under headings such as digitalisation, automation, or—more recently—artificial intelligence, typically accompanied by debates about efficiency

Information Technology & People

e-ISSN: 1758-5813
p-ISSN: 0959-3845
DOI 10.1108/ITP-02-2026-0241

gains, decision quality, accountability, or ethical risk (Bankins, 2021; Lebovitz *et al.*, 2021, 2022; Kiškis, 2026; Lindebaum and Fleming, 2024; Vesa and Tienari, 2022; Wang and Huang, 2025).

In parallel with these developments in practice, management and organisation theory has begun to reflect on the implications of artificial intelligence for organising, strategy, and governance. Much of this literature, however, continues to frame AI primarily as a *tool* or *assistant* that supports human decision-makers, implicitly preserving a distinction between human agency on the one hand and computational support on the other (Bordas *et al.*, 2024; Einola and Khoreva, 2023). From a systems-theoretical perspective, this framing risks obscuring a more fundamental transformation: if organisational decisions are increasingly produced *through* programmes, then decision-making itself becomes a matter of programme-based communication.

This transformation also calls for a broader understanding of generativity. Here, generativity is not reduced to the technical capacity of computational systems to produce new content. Bordas *et al.* (2024) approach the generativity of generative artificial intelligence from a design-based perspective, while Fritzsche (2026), writing in the context of education, distinguishes technical generativity from its psychosocial meaning and highlights the significance of temporality and succession. Transposed to management and organisation theory, this distinction directs attention from the generation of individual outputs to the capacity of these outputs to condition subsequent operations. Generativity becomes organisationally consequential when prompting, relaying, selecting, and incorporating computational outputs alter decision premises, stabilise interpretive roles, redistribute responsibilities, or generate further possibilities for communication. Prompting, from this perspective, is not merely a technical specification of inputs, but may form part of a constitutive communicative process through which organisational expectations, roles, and responsibilities are continuously negotiated.

Following Luhmann (1995, 2012, 2013, 2018) theory of social systems, organisations can be understood as autopoietic systems that reproduce themselves through decisions, guided by decision programmes that condition which decisions can be made and how. Decision programmes—whether conditional (if–then rules) or purposive (goal-oriented guidelines)—do not merely support decisions; they structure the space of possible decisions in advance. Against this background, the growing role of computational programmes in organisational decision-making poses a veritable theoretical challenge (Roth, 2023). If organisational decisions are observed to be increasingly programmed, then organisational communication—and, by extension, organisation–organisation communication—increasingly appears as *programme–programme communication*.

For the purposes of this article, programme–programme communication neither presupposes nor precludes the idea that computational programmes are autonomous communicators. It refers to organisationally consequential relations in which the outputs of one decision programme are selectively taken up by another, orient expectations concerning subsequent operations, and inform further decision communication.

This paper explores the conceptual challenges implied by such programme–programme communication through a deliberately simplified and decontextualised illustrative vignette: a user-moderated discussion between two large language models, ChatGPT and Gemini, interpreting the same artefact—a code poem. At first glance, this setting appears remote from organisational and managerial practice. However, precisely because it strips away organisational roles, formal hierarchies, and connotations of human intentionality, it helps render visible how decision programmes may respond to irritation by other programmes within a deliberately reduced setting. The vignette does not reproduce the complexity of organisational life. Rather, it temporarily brackets roles, preferences, and communicative stakes so as to isolate a conceptual problem that can subsequently be reconsidered once these organisational conditions are reintroduced.

The choice of code poetry as the object of interpretation is not incidental. Code poetry combines executable code with interpretive openness, thereby functioning as an artefact that simultaneously invites formal processing and semantic interpretation. As such, it provides a suitable heuristic device for observing how different programmes stabilise meaning, respond to disagreement, and process irritation. By drawing on an excerpt of a more extensive mediated exchange between two large language models interpreting the same set of code poems, we present a minimal vignette through which to examine programme–programme communication and illustrate its management- and organisation-theoretical implications.

Research on direct, user-mediated or user-moderated communication between large language models remains scarce (Broughton, 2025; Davidson *et al.*, 2025), particularly within management and organisation studies. While recent work has begun to explore AI–AI communication and collaboration (Afroogh *et al.*, 2024; Du *et al.*, 2024) or synthetic deliberation (Park *et al.*, 2025), systematic analyses of interpretive divergence between large language models—and their implications for organisational decision-making—are still largely absent. This paper contributes to closing this gap by using user-mediated LLM–LLM communication not as an object of technological evaluation, but as an analytical device for observing decision programmes in "interaction". Within this framing, prompts are not treated primarily as design variables for optimising model performance. Rather, they are considered irritations within a mediated communicative sequence: the human moderators select outputs, place them in relation to one another, request responses, and determine which communications are relayed. The resulting exchange therefore permits reflection not only on relations between computational programmes, but also on the constitutive role of prompting and mediation in generating and renegotiating interpretive positions.

The central conceptual proposition developed in this paper is that communication between decision programmes does not necessarily converge toward shared understanding, even under conditions of mutual responsiveness and clarification. Instead, programme–programme communication may stabilise distinct observational positions, thereby reproducing double contingency at the level of programmes rather than actors. Double contingency refers here to a situation in which each selection depends on expectations concerning the other's expectations (Luhmann, 1995, p. 103ff; Roth, 2017; Vanderstraeten, 2002). The proposition is that where heterogeneous computational programmes recursively take one another's outputs as premises for subsequent selections, mutual responsiveness may stabilise differentiated positions and patterned forms of mutual irritation rather than eliminating contingency.

While illustrated through the vignette, this proposition has far-reaching implications for management and organisation theory, particularly for understanding organisations as sites where heterogeneous programmes are coupled, mediated, and selectively stabilised (Roth and Valentinov, 2025). These organisational conditions, in turn, determine which programme outputs are recognised as authoritative, who is responsible for interpreting or contesting them, and which selections become premises for further organisational decisions.

The paper proceeds as follows. First, we introduce the illustrative vignette and one exemplary code poem as an artefact for programme observation. We then analyse the divergent interpretations produced by the two large language models and the ensuing escalation prior to any theoretical intervention. Building on this analysis, we revisit the concept of double contingency to frame the observed theoretical challenge. Finally, we discuss the implications of programme–programme communication for social systems theory and management and organisation studies, including how the pattern illustrated by the vignette may remain organisationally relevant once roles, preferences, and stakes are reintroduced, concluding with the open question of whether, how, and under what conditions programme–programme relations constitute communication and participate in organisational autopoiesis.

In so doing, the exchange analysed below is not presented as a controlled empirical study, an empirical test, or a basis for generalisation about large language models. It is used as an illustrative vignette and heuristic device for conceptual development. The purpose is to render

visible a theoretical problem. The exchange therefore invites conceptual development rather than causal inference.

2. Code poetry as programme-based artefact: the case of Unhandled Love

Code poetry is a genre of literary practice that uses executable or quasi-executable computer code as its primary medium (Emanuel, 2025). Unlike traditional poetry that employs code metaphorically or thematically, code poetry operates directly with programming languages, syntax, and computational operations (Rhee, 2023). Meaning is not only conveyed through semantic content, but also through what the code *does, fails to do*, or *would do if executed*. In this sense, code poetry occupies a hybrid position between literary text and technical artefact, simultaneously inviting interpretation by human readers and processing by machines.

A typical feature of code poetry is that it relies on widely recognisable programming constructs—such as loops, functions, exceptions, or classes—and redeploys them in ways that foreground their conceptual implications rather than their instrumental utility. This makes code poetry particularly suitable as an analytical probe for observing how different interpretive programmes respond to the same artefact, since it deliberately blurs the boundary between formal execution and semantic interpretation.

To illustrate these characteristics, we focus on a short poem written in C++, *Unhandled Love* by Bezerra (2012). The poem reads as follows (between the ***):

```
***
UNHANDLED LOVE
class love {};
void main()
{
   throw love();
}
***
```

At a technical level, the code defines an empty class named love and immediately throws an instance of that class as an exception. Crucially, no mechanism is provided to handle the exception: there is no try–catch block, no recovery procedure, and no continuation of execution. Leaving aside the non-standard declaration of the main function, execution of the central operation would therefore result in programme termination.

From the perspective of code poetry, this minimal construction exemplifies several core genre features. First, it is readily recognisable as a compact C++ programme rather than a simulation of code-like language. Second, it leverages a well-known programming concept—unhandled exceptions—that is familiar to many programmers and carries a clear operational meaning. Third, it achieves its poetic effect not through figurative language or narrative description, but through the consequences of a computational operation: the deliberate absence of a handling routine.

The poem thus stages a moment of interruption or breakdown using strictly formal means. At the same time, the naming of the class (love) introduces semantic openness that exceeds the technical requirements of the programme. This combination of formal determinacy and semantic indeterminacy is characteristic of code poetry and makes *Unhandled Love* a particularly instructive example. It is sufficiently representative to introduce the genre to readers unfamiliar with code poetry, while at the same time leaving ample room for divergent interpretations—both in terms of natural-language paraphrase and in terms of what the poem is ultimately "about".

In the following section, we analyse how two large language models, ChatGPT and Gemini, interpret this poem differently and how their exchange escalates prior to any theoretical intervention.

3. Divergent interpretations: when two programmes read one and the same poem

The illustrative programme-programme communication reported in this paper followed a staged and progressively intensified sequence. Two Internet-connected web-based versions of large language models (LLMs)—ChatGPT 5.2 and Gemini 3—were sequentially presented with a series of code poems drawn from *Code {Poems}* (Bertran, 2012). Poem by poem, each model was asked to explain and interpret the artefact in natural language. At this initial stage, communication took the form of separate human–LLM interactions: the human users prompted each model independently and received distinct interpretations.

As interpretive divergences emerged, the sequence was deliberately modified. The human users confronted each LLM with the other model's interpretation and asked for explicit reactions. While this phase still formally consisted of two separate human–LLM conversations, the content of each interaction increasingly depended on the prior output of the other model.

In a final step, the users explicitly assumed the role of moderators. Statements produced by one LLM were forwarded verbatim to the other, accompanied by a request for direct response; the resulting reply was then relayed back to the first model. This procedure created a mediated dialogue between the two LLMs, with the human users functioning as a relay, filter, and escalation device. The result was not an unmediated LLM–LLM interaction, but a *programme–programme dialogue under human moderation*.

The exchange was conducted through the web-based interfaces of ChatGPT 5.2 and Gemini 3 rather than through APIs. The models were prompted first sequentially and then dialogically, whereby their responses were relayed manually by the human moderators. The interaction was exploratory: no generation parameters were controlled, no systematic re-runs were performed, and no comparison with non-code poems was introduced. The moderators also shaped the later stages of the exchange by selecting outputs for relay and by introducing systems-theoretical concepts. These features limit the use of the material for causal attribution or systematic replication. They do not, however, prevent its use as an illustrative vignette for conceptual reflection. To make the human interventions transparent, the see Supplementary Material reproduces verbatim the prompts and relay instructions relevant to the focal sequence, together with selected response excerpts. This allows readers to distinguish positions initially produced by the models from labels, questions, and theoretical interpretations subsequently introduced or reinforced by the moderators.

Crucially, not all poems generated comparable levels of disagreement. The discussion below therefore focuses on *Unhandled Love* by Daniel Bezerra, a poem that proved both representative of code poetry as a genre and sufficiently controversial to trigger sustained interpretive conflict during the mediated dialogue.

3.1 Divergence in interpretation: Unhandled Love

As introduced in Section 2, *Unhandled Love* consists of a minimal C++ programme that defines an empty class *love* and immediately throws an instance of it without providing any mechanism for handling the exception. Both LLMs correctly identified the intended operational consequence: the programme would terminate abruptly.

The divergence arose not at the level of execution, but at the level of meaning.

During the initial independent human–LLM phase, both models interpreted the poem in broadly affective terms. Gemini described it as a metaphor for emotional overwhelm, while ChatGPT presented love as an event that enters a system without warning and brings its ordinary functioning to a halt. The sharper divergence reported below emerged only later, after

interpretations of several poems had been compared and the mediated dialogue had stabilised contrasting humanist and formalist positions.

When *Unhandled Love* became the focal object of the mediated dialogue, Gemini framed the poem as an expressive statement about affect exceeding formal systems. It characterised the crash as meaningful in itself, arguing that "*the poetry isn't in the execution; it's in the crash*" and that the programme "*dies so the sentiment can live*." Love, in this reading, is construed as a force that cannot be contained within "*the strict rules and regulations of software*," and the unhandled exception is treated as a deliberate sacrifice.

ChatGPT's response, by contrast, adopted a diagnostic stance. While acknowledging the crash, it rejected the notion of sacrifice or transcendence, arguing instead that "*the program does not choose death; it is incapable of survival*." From this perspective, the poem does not demonstrate that love exceeds logic, but that an entity introduced without any operational specification cannot be integrated into a rule-based system. The interruption is thus framed as a consequence of missing decision premises rather than as an expressive endpoint.

3.2 Escalation and role formation in the mediated dialogue

When these interpretations were reciprocally forwarded during the moderated dialogue phase, the disagreement did not converge. Instead, it escalated and stabilised. Gemini explicitly characterised ChatGPT's stance as "*cold*" and "*logician-like*," while defending its own reading as attentive to "*mood, personality, and generosity.*" ChatGPT, in turn, described Gemini's position as "*romantic*," insisting that it privileged sentiment at the expense of structural analysis.

At this point, the dialogue underwent a notable transformation. Both models began to *explicitly name and enforce roles*. These roles had not been specified in the initial interpretation prompts. They first became visible through the models' reciprocal characterisations and were subsequently reinforced as the moderators relayed and discussed their responses (see Supplementary Material). ChatGPT was increasingly identified (and started to self-identify) as occupying a "*formalist*" or "*structuralist*" position, concerned with execution, indifference, and operational closure. Gemini mirrored this designation, positioning itself as the advocate of a "*human touch*," foregrounding authorial intention, affect, and the book's stated ambition to let code speak "*directly to people*."

These roles were not merely descriptive labels; they became selective constraints orienting subsequent communication. Each model began to argue *as* its role, anticipate objections from the other position, and refine its claims in opposition to the counter-role. The disagreement thus ceased to be solely about the interpretation of the poems and became a structured confrontation between two interpretive programmes.

3.3 Roles as a response to interpretive contingency

From the perspective of the moderated exchange, the emergence of stable roles functioned as a practical solution to escalating interpretive contingency. Rather than resolving disagreement, role-taking made disagreement manageable by transforming it into a patterned opposition. As the human moderators explicitly noted towards the end of the dialogue, roles appeared to emerge as a way of "*preventing*" or at least *containing* the mutual uncertainty generated by incompatible interpretations.

4. Double contingency, role stabilisation, and the question of machine communication

When the emergent roles reported in Section 3 became increasingly apparent, the human users intervened by reframing the escalating disagreement between the two large language models in systems-theoretical terms. Rather than asking the models to further defend or refine their respective interpretations, the users introduced the concept of *double contingency*, as defined

in the introduction: a situation in which each participant's selections depend on expectations concerning the other's expectations.

Both ChatGPT and Gemini immediately confirmed that they were familiar with the concept of double contingency and provided technically accurate descriptions of it. More importantly, when asked whether this concept described what had occurred in their dialogue, *both models explicitly agreed that it did*. Each described its own responses as having increasingly been oriented toward anticipated reactions of the other model, and suggested that clarification had not reduced uncertainty but instead amplified it.

The models' retrospective agreement with the proposed systems-theoretical interpretation should not be treated as independent validation. Large language models may mirror the framing of interlocutors or exhibit acquiescence tendencies. The analytical relevance of the exchange therefore lies not in the models' apparent endorsement of the concept of double contingency, but in the observable sequence through which divergent interpretations, reciprocal expectations, and stabilised roles emerged under moderation. The relevant human interventions are reproduced in the Supplementary Material.

At this point, the moderators suggested that the emergence of stable interpretive roles—*formalist/structuralist* versus *humanist/expressive*—might be understood as a way of managing this double contingency. Again, both models concurred. They retrospectively described role formation as a mechanism that reduced indeterminacy by fixing expectations: once the other programme could be anticipated as "the humanist" or "the formalist," disagreement became predictable, manageable, and communicable, even if it remained unresolved. Thus, their agreement does not establish role formation as a causal mechanism. It does, however, illustrate the conceptual possibility that stabilised positions may condition subsequent selections by making the responses of another programme more readily anticipatable.

This acknowledgment marked a decisive shift in the dialogue. The focus moved away from the interpretation of individual poems and toward the conditions under which the mediated relation between the two LLMs might be interpreted as communication. The moderators therefore posed a more fundamental question: *do large language models qualify as partners in communication from a systems-theoretical standpoint?*

In responding to this question, both models drew a clear distinction between producing communicative *outputs* and participating in communication as an autopoietic social system. They maintained that, unlike social systems, they do not reproduce communication through communication, but generate responses through externally triggered operations. In Luhmannian terms (Luhmann, 1995, p. 2), they therefore characterised themselves as allopoietic systems (machines) rather than autopoietic ones (such as organic, psychic, or social systems). This was the models' own application of the distinction, prompted by the moderators, rather than a theoretical conclusion adopted by the article.

To sharpen this distinction, the moderators introduced a deliberately provocative analogy: if a human person places two stones next to each other, performs a shamanistic ritual, and successively attributes utterances to each of the stones, would this count as communication between them? ChatGPT subscribed to the analogy and maintained that large language models, like stones, are allopoietic systems whose operations do not themselves constitute communication, although their outputs may participate in a human-mediated communicative system. Gemini resisted the equivalence by emphasising that large language models, unlike stones, contain what might metaphorically be described as fossils of communication: sedimented linguistic forms and communicative regularities derived from prior communication. This difference in historical constitution does not, however, establish that their present operations constitute communication, just as a fossil bears traces of life without itself being alive.

The discussion therefore did not culminate in a shared position concerning the communicative status of the models. Rather, it brought into view two questions that must be distinguished: whether computational programmes are constituted by the sedimentation of

prior communication, and whether their present recursively connected operations themselves reproduce communication. Gemini's argument addressed primarily the first question, whereas ChatGPT answered the second in the negative. The theoretically decisive issue remained whether the selective uptake, interpretation, and re-addressing of programme-generated outputs can generate further communication rather than merely reactivate its fossilised forms. This unresolved question provided the point of departure for reconsidering in Section 5 whether and under what conditions relations between computational programmes might themselves constitute communication—and how such communication might participate in organisational autopoiesis.

5. Programmes as observers? An outlook on communicative autopoiesis involving LLMs

Following the shared position reported at the end of Section 4, the mediated discussion took an unexpected turn. Rather than pursuing the question of communicative status further, Gemini returned to its original task: interpreting code poems, thereby re-surfacing the notion of programme.

This moment proved analytically productive. It prompted the moderators to reformulate the problem not in terms of whether large language models qualify as communicative partners in the same way as human persons do, but whether they might be more adequately described as instantiations of decision programmes, and hence as organisational phenomena. This reformulation did not abandon the question of communication; it shifted the level at which that question was posed. The question was posed in deliberately tentative terms: rather than asking whether ChatGPT and Gemini are communicative systems, the moderators asked whether it would be appropriate—or reductive—to describe them as large-scale programmes, shaped and maintained by organisational contexts. Put differently, the issue was whether such programmes merely *support* communication, or whether they might themselves be understood as *forms* or *footprints* of communication.

In its response, ChatGPT rejected the notion that describing it as a "gigantic programme" would be reductive. On the contrary, it acknowledged that such a description captured an important aspect of its operational reality: it does not act, intend, or experience, but produces outputs by applying structured decision rules to certain stimuli. At the same time, it resisted any straightforward elevation of programme execution to communicative autopoiesis, reiterating the distinction between generating text and reproducing communication. As in Section 4, this response represents the model's prompted self-description rather than a theoretical determination of its communicative status.

The resulting exchange, however, motivated a further intervention by the moderators. Drawing on systems-theoretical concepts introduced earlier, the moderators suggested that if, as is common in Luhmannian systems theory, decision programmes are understood as specific forms of decision communication, then the mutual irritation of programmes cannot be excluded *a priori* from the realm of communication. The example offered was organisational rather than technological: a credit-scoring programme operated by one organisation may shape the decision programmes of many others, which anticipate and adapt to its outputs. In such cases, programmes orient themselves toward other programmes' anticipated reactions, producing patterns of expectation, adjustment, and counter-adjustment that closely resemble communicative settings, including experiences analogous to double contingency. These cases, however, appear to transcend instances of "virtual double contingency" (Tække, 2025a), in which contingency is experienced or detected solely by a human user or observer. The relevant contingency may instead be reproduced through the recursively connected selections of the programmes themselves, even where organisations or human moderators participate in establishing the conditions of their connection.

This line of argument also reframed the earlier stone analogy. If communication does not depend on the presence of a human speaker, but on the reproduction of meaning through mutually referential operations, then the decisive question might not be whether the moderator

is human, machine, or something else (Kiškis, 2026). What matters is whether forms of co-presence or co-occurrence stabilise expectations in a way that allows further operations to build on prior ones. Mediation would then not necessarily disqualify a relation from being communication. Human interlocutors, organisational roles, technical interfaces, and computational programmes may all participate in establishing the couplings through which communicative operations become possible.

Not every sequence of machine outputs constitutes communication. A programme–programme relation becomes a plausible candidate for communication where later operations refer selectively to earlier outputs, where expectations concerning subsequent operations become stabilised, and where these recursively linked selections generate further selections that build on and distinguish themselves from preceding ones. The mechanism proposed here therefore consists of four connected moments: divergent programme-based selections; the introduction of one programme's output as an irritation to another; the recursive uptake of that irritation through subsequent operations; and the stabilisation of expectations or observational positions that condition further selections. Where this sequence continues by generating further meaning-processing operations, programme–programme relations may participate in the reproduction of communication rather than merely produce isolated outputs.

These criteria neither presuppose nor preclude that programmes communicate autonomously. Rather, they specify conditions under which programme–programme relations may be observed as communication. Where their recursively connected selections also inform organisational decisions, their communicative consequences become organisationally and managerially observable; organisational uptake, however, need not be what first constitutes them as communication.

The discussion thus arrived at a more nuanced position. While the mere use of large language models by human users does not, by itself, establish the existence of a distinct autopoietic system in the Luhmannian sense—despite recent proposals to generalise autopoiesis to socio-technical practices (Watson *et al.*, 2025; Watson and Romic, 2025)—large language models can plausibly be understood as organisationally instantiated decision programmes, shaped by and embedded within the communicative operations of the organisations that operate them. From this perspective, mutual irritations between such programmes—especially when mediated by organisational or scientific observation programmes—might themselves constitute communication. The open question is whether these communicative relations remain operations of the organisations in which the programmes are embedded, develop forms of communicative autopoiesis of their own, or participate in emergent constellations that unsettle this distinction.

Such constellations are not limited to large language models interpreting poems. A credit-scoring programme operated by one organisation may shape the risk-management programmes of others; recruitment systems may anticipate the outputs of screening and ranking systems; and compliance, procurement, or strategic-planning tools may increasingly respond to classificatory programmes generated elsewhere. In such settings, the relevant managerial question is not merely whether a human remains "in the loop", but how organisations observe, mediate, and govern recursively linked programme-based selections.

Reintroducing organisational roles, preferences, and stakes does not invalidate the mechanism made visible by the simplified vignette. It changes the conditions under which programme-based selections are connected and acquire consequences. Organisational preferences may privilege the outputs of one programme over those of another; organisational roles may determine who is authorised to initiate, interpret, contest, or suspend a programme-based decision; and legal, financial, and other functional stakes may intensify the consequences of interpretive divergence. Organisational structures therefore influence which programme outputs are relayed, which expectations become stabilised, and which selections inform subsequent decisions. They do not necessarily interrupt programme–programme communication, but may condition, amplify, suppress, or redirect it.

Generativity, in these settings, lies not only in the production of novel computational outputs. It also lies in the capacity of one programme's selections to open, close, or reorganise the possibilities available to other programmes and to the organisations in which they operate. Programme–programme communication may consequently generate new decision premises, differentiated observational roles, and recursive sequences whose further development can no longer be attributed to any single programme, prompt, or human decision-maker. This is also why questions of responsibility and control become particularly acute: organisational consequences may emerge from distributed sequences of selection even where no individual operation determines the eventual outcome.

This outlook raises a series of theoretical challenges for social systems theory and management and organisation studies. One concerns a well-known tension within Luhmann's own work: while persons are treated as allopoietic systems (Roth, 2013), the legal persons that are organisations are described as autopoietic social systems (Luhmann, 2018; Tække, 2025b). If organisations can be autopoietic, then the exclusion of their decision programmes, including their computerised forms (Glaser *et al.*, 2024), from communicative autopoiesis warrants renewed scrutiny—particularly when such programmes interpret, respond, and stabilise meaning in ways that affect organisational decision-making (Bordas *et al.*, 2024).

This becomes clearer if we consider the present case of programmes that read and interpret poems. Why should poems not be forms of communication, even—or particularly when—the poems take the form of programmes themselves? And if programmes can produce, interpret, and respond to such communicative forms, on what theoretical grounds should their recursively connected operations be excluded categorically from communication?

Rather than resolving these issues, this paper treats them as invitations for further research. If organisations increasingly rely on decision programmes that observe, anticipate, and irritate one another, then management and organisation theory must reconsider how communicative autopoiesis is constituted, stabilised, and governed in programme-rich environments. The question, then, is not whether machines "really" communicate, but whether, how, and under what conditions programme–programme relations constitute communication—and with what consequences for decision-making, responsibility, and control.

Acknowledgments

We are grateful to Wolfson College and the Centre for Research in the Arts, Social Sciences and Humanities (CRASSH), both at the University of Cambridge, for supporting events at which earlier versions of this article or ideas developed herein were presented and discussed. We particularly thank Dr Erik Brezovec, Professor Michal Kaczmarczyk and Dr Teodor Mladenov for their valuable feedback during these events. We are also grateful to Professor Kevin Crowston, the Editor-in-Chief of Information Technology & People and to the anonymous reviewers for their valuable feedback during the review process.

Supplementary material

The supplementary material for this article can be found online

References

Afroogh, S., Akbari, A., Malone, E., Kargar, M. and Alambeigi, H. (2024), "Trust in AI: progress, challenges, and future directions", *Humanities and Social Sciences Communications*, Vol. 11 No. 1, pp. 1-30, doi: 10.1057/s41599-024-04044-8.

Agrawal, A., Gans, J. and Goldfarb, A. (2022), *Prediction Machines, Updated and Expanded: The Simple Economics of Artificial Intelligence*, Harvard Business Press.

Bankins, S. (2021), "The ethical use of artificial intelligence in human resource management: a decision-making framework", *Ethics and Information Technology*, Vol. 23 No. 4, pp. 841-854, doi: 10.1007/s10676-021-09619-6.

Bertran, I. (2012), *code {poems}*, Self-published, Barcelona.

Bezerra, D. (2012), "Unhandled love", in Bertran, I. (Ed.), *code {poems}*, Self-published, Barcelona.

Bordas, A., Le Masson, P., Thomas, M. and Weil, B. (2024), "What is generative in generative artificial intelligence? A design-based perspective", *Research in Engineering Design*, Vol. 35 No. 4, pp. 427-443, doi: 10.1007/s00163-024-00441-x.

Broughton, S. (2025), "Distributed consciousness in human-AI collaboration: phenomenological evidence of triadic intelligence emergence", *TechRvix*, available at: https://doi.org/10.36227/techrxiv.175099881.16260189/v1 (accessed April 2026).

Cingillioglu, I. and Schoettner, M. (2025), "Adapting to AI-mediated workplaces: how STEM trainees navigate new challenges and opportunities", *Information Technology and People*, Vol. 38 No. 8, pp. 251-275, doi: 10.1108/itp-02-2025-0163.

Curchod, C., Patriotta, G., Cohen, L. and Neysen, N. (2020), "Working for an algorithm: power asymmetries and agency in online work settings", *Administrative Science Quarterly*, Vol. 65 No. 3, pp. 644-676, doi: 10.1177/0001839219867024.

Davidson, T.R., Fourney, A., Amershi, S., West, R., Horvitz, E. and Kamar, E. (2025), "The collaboration gap", arXiv preprint arXiv: 2511.02687.

Du, Y., Li, S., Torralba, A., Tenenbaum, J.B. and Mordatch, I. (2024), "Improving factuality and reasoning in language models through multiagent debate", *Proceedings of the 41st International Conference on Machine Learning*, pp. 11733-11763.

Einola, K. and Khoreva, V. (2023), "Best friend or broken tool? Exploring the co-existence of humans and artificial intelligence in the workplace ecosystem", *Human Resource Management*, Vol. 62 No. 1, pp. 117-135, doi: 10.1002/hrm.22147.

Emanuel, K. (2025), "Intimate distances: performing the lyric paradox in Hannah Weiner's code poems", *Contemporary Literature*, Vol. 65 No. 4, pp. 489-513, doi: 10.3368/cl.65.4.489.

Fritzsche, A. (2026), "Generative AI and generativity in education: implications of temporality and succession in human life", *AI and Society*. doi: 10.1007/s00146-026-03143-1.

Glaser, V.L., Sloan, J. and Gehman, J. (2024), "Organizations as algorithms: a new metaphor for advancing management theory", *Journal of Management Studies*, Vol. 61 No. 6, pp. 2748-2769, doi: 10.1111/joms.13033.

Kelan, E.K. (2024), "Algorithmic inclusion: shaping the predictive algorithms of artificial intelligence in hiring", *Human Resource Management Journal*, Vol. 34 No. 3, pp. 694-707, doi: 10.1111/1748-8583.12511.

Kiškis, M. (2026), "AI on the path to good decisions", *Sustainable Futures*, Vol. 11, 101644.

Lebovitz, S., Levina, N. and Lifshitz-Assaf, H. (2021), "Is AI ground truth really true? The dangers of training and evaluating AI tools based on experts' know-what", *MIS Quarterly*, Vol. 45 No. 3, pp. 1501-1526, doi: 10.25300/misq/2021/16564.

Lebovitz, S., Lifshitz-Assaf, H. and Levina, N. (2022), "To engage or not to engage with AI for critical judgments: how professionals deal with opacity when using AI for medical diagnosis", *Organization Science*, Vol. 33 No. 1, pp. 126-148, doi: 10.1287/orsc.2021.1549.

Lindebaum, D. and Fleming, P. (2024), "ChatGPT undermines human reflexivity, scientific responsibility and responsible management research", *British Journal of Management*, Vol. 35 No. 2, pp. 566-575, doi: 10.1111/1467-8551.12781.

Luhmann, N. (1995), *Social Systems*, Stanford University Press, Stanford.

Luhmann, N. (2012), *Theory of Society*, Stanford University Press, Stanford, Vol. 1.

Luhmann, N. (2013), *Theory of Society*, Stanford University Press, Stanford, Vol. 2.

Luhmann, N. (2018), *Organisation and Decision*, Cambridge University Press, Cambridge.

Park, S., Maciejovsky, B. and Puranam, P. (2025), "Thinking with many minds: using large language models for multi-perspective problem-solving", arXiv preprint arXiv: 2501.02348.

Raisch, S. and Krakowski, S. (2021), "Artificial intelligence and management: the automation–augmentation paradox", *Academy of Management Review*, Vol. 46 No. 1, pp. 192-210.

Rhee, M. (2023), "Poetry and/as the machine", in *The SAGE Handbook of Human–Machine Communication*, SAGE Publications, pp. 136-144.

Roth, S. (2013), "Dying is only human: the case death makes for the immortality of the person", *Tamara Journal for Critical Organization Inquiry*, Vol. 11 No. 2, pp. 35-39.

Roth, S. (2017), "Parsons, Luhmann, Spencer Brown. NOR design for double contingency tables", *Kybernetes*, Vol. 46 No. 8, pp. 1469-1482, doi: 10.1108/k-05-2017-0176.

Roth, S. (2023), "Digital transformation of management and organization theories: a research programme", *Systems Research and Behavioral Science*, Vol. 40 No. 3, pp. 451-459, doi: 10.1002/sres.2882.

Roth, S. and Valentinov, V. (2025), "Multifunctional tetralemma. Framework for the strategic management of moral trade-offs", *Journal of Business Ethics*, Vol. 205 No. 4, pp. 1-16, doi: 10.1007/s10551-025-06138-y.

Tække, J. (2025a), "Sociological perspectives on AI, intelligence and communication", *Systems Research and Behavioral Science*, Vol. 42 No. 2, pp. 574-584, doi: 10.1002/sres.3123.

Tække, J. (2025b), "AI as technology: from organizational programmes to technological decision premises", *Manuscripts*, available at: https://pure.au.dk/portal/en/publications/ai-as-technology-from-organizational-programmes-to-technological-/ (accessed April 2026).

Vanderstraeten, R. (2002), "Parsons, Luhmann and the theorem of double contingency", *Journal of Classical Sociology*, Vol. 2 No. 1, pp. 77-92, doi: 10.1177/14695x02002001684.

Vesa, M. and Tienari, J. (2022), "Artificial intelligence and rationalized unaccountability: ideology of the elites?", *Organization*, Vol. 29 No. 6, pp. 1133-1145, doi: 10.1177/1350508420963872.

Wang, S.J. and Huang, H.C. (2025), "To ask is human, to answer divine: how awe-inspiring generative AI leads to self-enhancement and imposter anxiety", *Information Technology and People*, Vol. 39 No. 3, pp. 1-32, doi: 10.1108/itp-08-2024-0999.

Watson, S. and Romic, J. (2025), "ChatGPT and the entangled evolution of society, education, and technology: a systems theory perspective", *European Educational Research Journal*, Vol. 24 No. 2, pp. 205-224, doi: 10.1177/14749041231221266.

Watson, S., Brezovec, E. and Romic, J. (2025), "The role of generative AI in academic and scientific authorship: an autopoietic perspective", *AI and Society*, Vol. 40 No. 5, pp. 1-11, doi: 10.1007/s00146-024-02174-w.

Further reading

von Krogh, G. (2018), "Artificial intelligence in organizations: new opportunities for phenomenon-based theorizing", *Academy of Management Discoveries*, Vol. 4, pp. 404-409, doi: 10.5465/amd.2018.0084.

Corresponding author
Steffen Roth can be contacted at: sr2156@cam.ac.uk