



METHODOLOGY-GOVERNED AI FOR REGENERATIVE URBAN PLANNING FROM SINGLE-PASS TO ROLE-DECOMPOSED SCM EXECUTION IN DE MARKEN, ALMERE

DR. MARINA MATASHOVA¹ (MARINA@ANDORRA-LAB.COM)¹

ELENA SELEZNEVA² (EVSELEZNEVA@ALMERE.NL)²

¹Co-FOUNDER, ANDORRA-LAB, PROFESSOR, UNIVERSITY OF BARCELONA IL-3

²GEMEENTE ALMERE/MUNICIPALITY OF ALMERE, URBAN PLANNER

KEYWORDS	ABSTRACT
Methodology-Governed AI; Role-Decomposed Workflow; Generative AI; Public-Realm Renewal; Regenerative Urban Planning; Sustainable Community Masterplan (SCM); Municipal Decision Support.	<i>This study tests methodology-governed LLM execution for regenerative public-realm planning in De Marken, Almere. Under a frozen municipal corpus, it compares a structured non-SCM baseline (A), single-pass Sustainable Community Masterplan (SCM) execution (B), and a staged, role-decomposed SCM workflow (C). Explicit SCM specification strengthened relational territorial reasoning and regenerative direction. Staged execution better preserved that direction through portfolio re-validation, delivery structuring and rendering. Research scores ranked C>B>A; a blind municipal evaluation ranked C>A>B. Methodological gains do not automatically become decision usefulness: overall usefulness stayed low in both single-pass conditions and rose only under staged execution. The study distinguishes specification from execution architecture and treats methodology-governed AI as municipal decision support, not autonomous planning.</i>

RECEIVED: 20/ 06 / 2026

ACCEPTED: 10/ 08 / 2026

1. Introduction

Interpretar ya es proponer
Manuel de Solà-Morales (1981)

Regenerative urban planning does not begin by adding climate, social and technical measures to a conventional renewal list. It begins with interpretation: reading the territory as a living socio-ecological system and deciding which relationships to protect, repair, intensify or leave low-intensity (Reed, 2007; Mang & Reed, 2012; Cole et al., 2013).

De Marken makes that requirement concrete. It is a *bloemkoolwijk* in Almere Haven whose nested hofjes, woonerven, green-blue structure and access routes already form one public-realm system. The forthcoming *Groot Onderhoud* (GOW) programme is legally framed as maintenance of existing public space. The live question is whether a place-based regenerative reading can enter that municipal package without being replaced by a fluent catalogue of works.

Municipal briefs typically arrive as separated assets, deficits and responsibilities. That packaging is not a defect of the brief. It is the working infrastructure of renewal. A regenerative reading has to be carried through it on purpose.

Large language models can already synthesise heterogeneous planning material and generate plausible propositions (Pan et al., 2025; Fu et al., 2025; Zheng et al., 2025). The difficulty is not generation. It is whether place-based relational reasoning survives the move from interpretation to a decision-facing package (Son et al., 2023). With LLMs a methodology can be supplied as knowledge, or operationalised as an executable specification: what may count as evidence, which transitions are required, what may be translated for a decision audience, and which commitments the specification may not close. The second form is treated here as governance through design (Shneiderman, 2020; Hong et al., 2024).

The study tests that proposition through the Sustainable Community Masterplan (SCM), a practice-based methodology for regenerative urban planning (Matashova, 2021, 2022). SCM is operationalised as a shared governance specification and compared across three controlled configurations:

- A. structured non-SCM baseline
- B. SCM-encoded single-pass execution
- C. role-decomposed SCM workflow

The comparison asks whether explicit specification changes regenerative reasoning, and whether staged execution changes what is preserved in the municipal package. Human authority was not experimentally manipulated. It remains a boundary condition of decision support rather than a substitute for municipal judgement.

1.1. Research gap

The study distinguishes methodology as knowledge supplied to a model from methodology in operational use. The first asks whether professional concepts, relations and operations are made explicit. The second asks whether those commitments survive execution: from territorial interpretation to a municipal decision-support output.

Regenerative theory specifies the reasoning required: relational, place-based and co-evolutionary (Reed, 2007; Mang & Reed, 2012; du Plessis, 2012; Cole et al., 2013; Cole, 2023). SCM gives that requirement an operational sequence (Matashova, 2021, 2022). This literature defines the object of the experiment. It does not specify how the sequence is operationalised when a language model is the working medium.

AI-supported professional workflows can externalise domain knowledge through task representations, intermediate products and specialised roles (Wong et al., 2024; Hong et al.,

2024; Grand et al., 2025; Wang et al., 2024). Planning applications likewise combine theoretical constraints, evidence grounding and staged workflows (Wang, Fu & Lyles, 2026; Huang et al., 2026; Zhu, Chen & Zhang, 2025). Less clear is how this principle extends from structuring tasks to a methodology whose function is also to structure how a territory is interpreted and then translated into strategy and decision support.

The unresolved question is therefore not whether professional knowledge can be operationalised for LLM execution, but whether methodological commitments are preserved through successive transformations towards a municipal package.

Human–AI research treats the distribution of control as a design problem (Amershi et al., 2019; Li et al., 2023). In this study, human authority is a boundary condition of decision support, not an experimentally manipulated variable. The workflow is designed to support situated professional judgement, not to displace municipal decision rights.

1.2 Research questions and objectives

Two research questions and one analytical boundary:

1. RQ1. What differences in regenerative-planning reasoning and municipal decision usefulness are observed when explicit SCM specification is introduced within otherwise controlled single-model, single-pass execution?
2. RQ2. How do methodological preservation and decision-facing translation differ between SCM-encoded single-pass execution and a role-decomposed SCM workflow?

The study compares three experimental conditions, summarised in Figure 1:

Condition A receives the municipal question, frozen corpus and common output specification, without SCM (Figure 1).

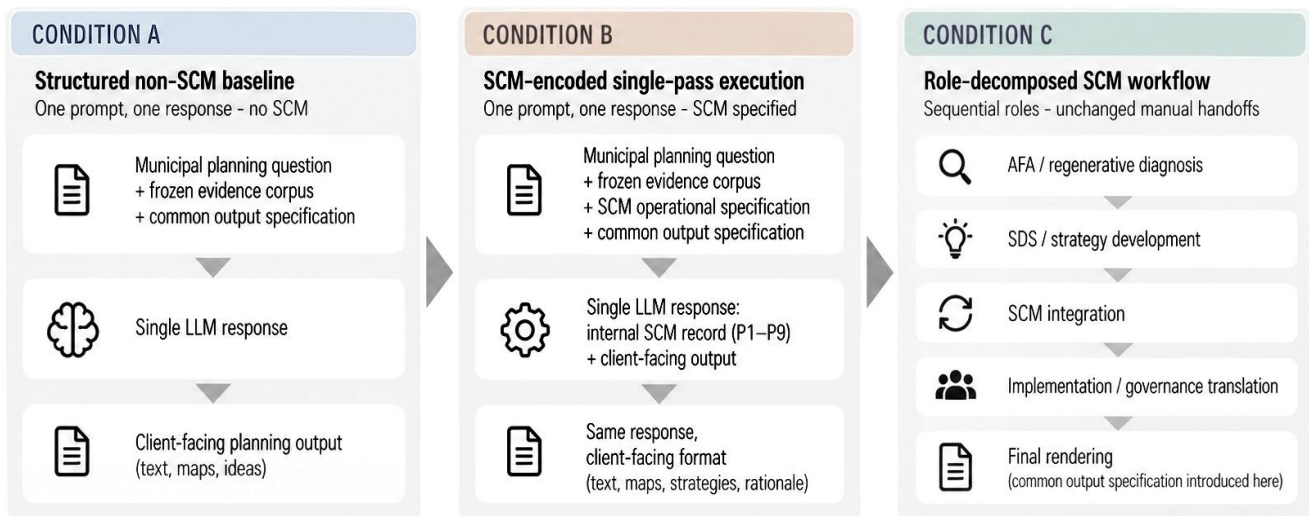
Condition B receives the same inputs plus the SCM specification, executed in one prompt, one continuous context and one pass.

Condition C executes the same SCM specification through bounded sequential roles with unchanged handoffs; the common output specification is introduced only at final rendering.

The A→B contrast isolates the introduction of SCM under single-pass execution. The B→C contrast compares two ways of executing the same methodological foundation and therefore bundles a change in execution configuration.

Analytical boundary. Human authority was not manipulated and remained outside both contrasts.

Figure 1. Experimental configurations A–C.



Source: Author 1, 2026.

1.3 Contribution of the study

The study contributes a controlled comparison of two ways of operationalising SCM — single-pass execution and a staged, role-decomposed workflow — against a structured non-SCM baseline in a live municipal case. Regenerative-planning reasoning and municipal decision usefulness are assessed as distinct constructs. Process tracing then follows how methodological commitments were preserved, transformed or lost.

Four distinctions follow and are developed in the discussion:

1. Specifying professional reasoning does not ensure its preservation through transformation.
2. Regenerative continuity depends on re-validating relationships against a place-based Pattern, not on retaining prior context alone.
3. Institutional delivery and client-facing rendering are downstream functions: they may operationalise or re-present the proposition, but they may not replace the Pattern or close what the evidence does not warrant.
4. Methodological integrity and municipal decision usefulness can diverge.

Together these distinctions describe how a regenerative reading of place can remain operative as it is turned into a municipal decision package. The tested configuration is positioned as decision support at TRL 5, not as autonomous planning.

2. Conceptual Framework: From Regenerative Planning Methodology to AI-Supported Reasoning

2.1. Regenerative urbanism as a whole-system reasoning challenge

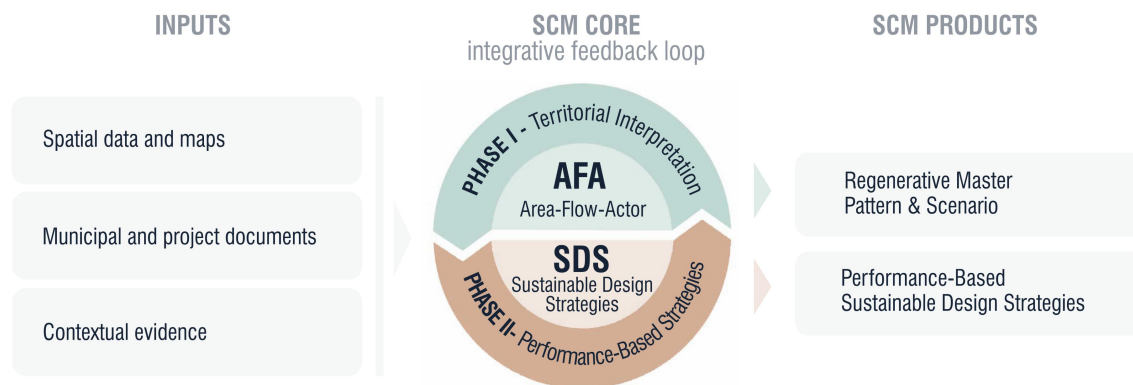
Regenerative development treats interventions as contributions to the health, adaptive capacity and continuing development of socio-ecological systems, rather than as the reduction of negative impacts alone (Reed, 2007; Mang & Reed, 2012; du Plessis, 2012; Cole et al., 2013; Cole, 2023). In urban planning, this requires interpreting place as a living socio-ecological system: reading relationships among spatial form, ecological processes, flows, everyday patterns of use, institutions and stewardship, and understanding how they shape both the present territorial condition and its capacity to evolve.

The study therefore treats **regenerative planning as a whole-system reasoning challenge**, not as a new test of regenerative theory. Conventional planning commonly responds to pre-given assets, problems and delivery packages. Additive sustainability may attach climate, biodiversity, accessibility and social objectives to that package. Regenerative planning instead re-reads the territory: which relationships organise the system, where they reinforce or constrain one another, where discontinuities occur, and how changing one relationship may alter the whole. Structural, performance and normative layers remain necessary, but only if this territorial logic remains available as reasoning moves from interpretation to strategy and decision support.

2.2. SCM as an operational framework for regenerative urban planning

The Sustainable Community Masterplan (SCM) gives this whole-system reasoning an operational structure rather than a catalogue of themes (Figure 2). Its methodological architecture combines two linked stages-AFA and SDS-with an iterative integration process connecting territorial interpretation to performance-based action.

Stage I - AFA territorial interpretation. Area-Flow-Actor analysis, together with the Two-Network reading of ecological and socio-technical systems, reconstructs the territorial condition as a relational whole. This stage is primarily qualitative and interpretive: it identifies relationships, dependencies, discontinuities, conflicts and latent capacities, and locates the place on a conventional-sustainable-regenerative transition. From this reading, SCM derives a Regenerative Master Pattern and spatial Scenario, establishing a place-

Figure 2. SCM reasoning architecture used in the experiment

Source: Author 1, 2026.

specific direction for what should be protected, repaired, intensified, reconnected or left low-intensity.

Stage II - SDS performance-based operationalisation. Sustainable Design Strategies translate the Regenerative Scenario into coordinated, multi-domain strategies, criteria and measurable performance requirements. This stage is primarily operational and accountable: it moves from territorial direction to concrete design, planning and governance strategies while making their intended effects testable. Metrics therefore do not independently define the regenerative direction; they provide a means of assessing whether strategies derived from the Scenario are producing the intended systemic effects.

Integration - iterative relational testing and feedback. Integration connects the interpretive and performance-based dimensions of SCM. It tests whether individual and cross-domain strategies actually reinforce the Pattern and Scenario, whether optimisation within one domain undermines another, and whether measurable gains remain connected to the territorial relationships that gave rise to the regenerative direction. The process is therefore iterative rather than additive: strategy and performance feedback can require relationships, synergies, trade-offs or implementation assumptions to be reconsidered before an integrated proposition is stabilised (Matashova, 2021, 2022).

2.3. Operationalising SCM for controlled LLM execution

For the experiment, SCM was formalised as a shared, case-independent Methodological Reference for Conditions B and C. To make the methodology executable and traceable, its reasoning was translated into a sequence of analytical products (P1–P9), shown in Table 1. These products preserve the progression from AFA interpretation through SDS strategy development and SCM integration, before the completed reasoning is translated towards municipal implementation and the common client-facing output.

SCM Methodological / translation layer	Function	Experimental product	
		Index	Content
SCM Stage I - AFA (Area-Flow-Actor)	Relational territorial interpretation through AFA and Two-Network analysis; establishment of regenerative territorial direction	P1	AFA Three-Axis Territorial Interpretation
		P2	Cross-Axis AFA Synthesis / Territorial Mental Map

		P3	Regenerative Master Pattern
		P4	Regenerative Scenario
SCM Stage II - SDS (Sustainable Design Strategies)	Translate the Regenerative Scenario into coordinated multi-domain strategies, criteria and measurable performance requirements.	P5	SDS Domain Baselines
		P6	SDS Strategy & Synergy Matrix
		P7	Cross-Domain Synergy Framework
SCM transversal integration	Re-test strategies and performance against Pattern and Scenario; identify synergies, trade-offs and domain optimisation.	P8	Integrated SCM Proposition
Institutional / municipal implementation translation	Translate the integrated SCM proposition into municipal delivery and institutional logic, including governance, responsibilities, sequencing and implementation conditions.	P9	Implementation & Governance Translation
Client-facing translation	Translate the completed analytical record into the Common Output Specification without introducing new substantive analysis.	Final output	Client-facing experimental deliverable

Table 1. Translation of the SCM sequence into traceable LLM execution products

Source: Author 1, 2026.

This common methodology-derived sequence provided the shared analytical structure for Conditions B and C; its implementation through different execution configurations is described in Section 4.

3. Case Study: De Marken, Almere

3.1. Territorial and planning context

De Marken is a residential neighbourhood in Almere Haven whose public realm expresses Dutch bloemkoolwijk morphology: clustered housing, courtyards, low-traffic streets and a continuous green-blue setting organised through a nested, branching system (Bijlsma et al., 2008; Rijksdienst voor het Cultureel Erfgoed, 2020) as shown in Figure 3.

Access routes feed woonerven, hofjes and cul-de-sac streets, linking movement, residential access, green space, facilities and everyday activity within one public-realm system. This inherited morphology combines recognised human-scale, low-traffic and green qualities with contemporary challenges of accessibility, public-space use, ecological performance and long-term management (Rijksdienst voor het Cultureel Erfgoed, 2020). It therefore provides a demanding case for integrated planning, in which spatial organisation, use, environmental performance, community value and maintenance are interdependent. Local analysis is grounded in De Marken evidence (E06–E14); published New Town research provides typological context only (de Meere et al., 2024).

3.2. Groot Onderhoud as the live context for integrated neighbourhood renewal

The planned Groot Onderhoud (GOW) programme provides the immediate planning context: renewal and long-term stewardship of existing public space, with municipal analysis also exploring links to climate adaptation, biodiversity, liveability, accessibility and social objectives (Gemeente Almere, 2026).

For the experiment, this context was reconstructed from the supplied municipal documents as a Municipal Brief framing GOW as a potential vehicle for moving from sectoral asset renewal towards an integrated, place-based public-realm approach.

Figure 3. De Marken study area and experimental context in Almere, the Netherlands

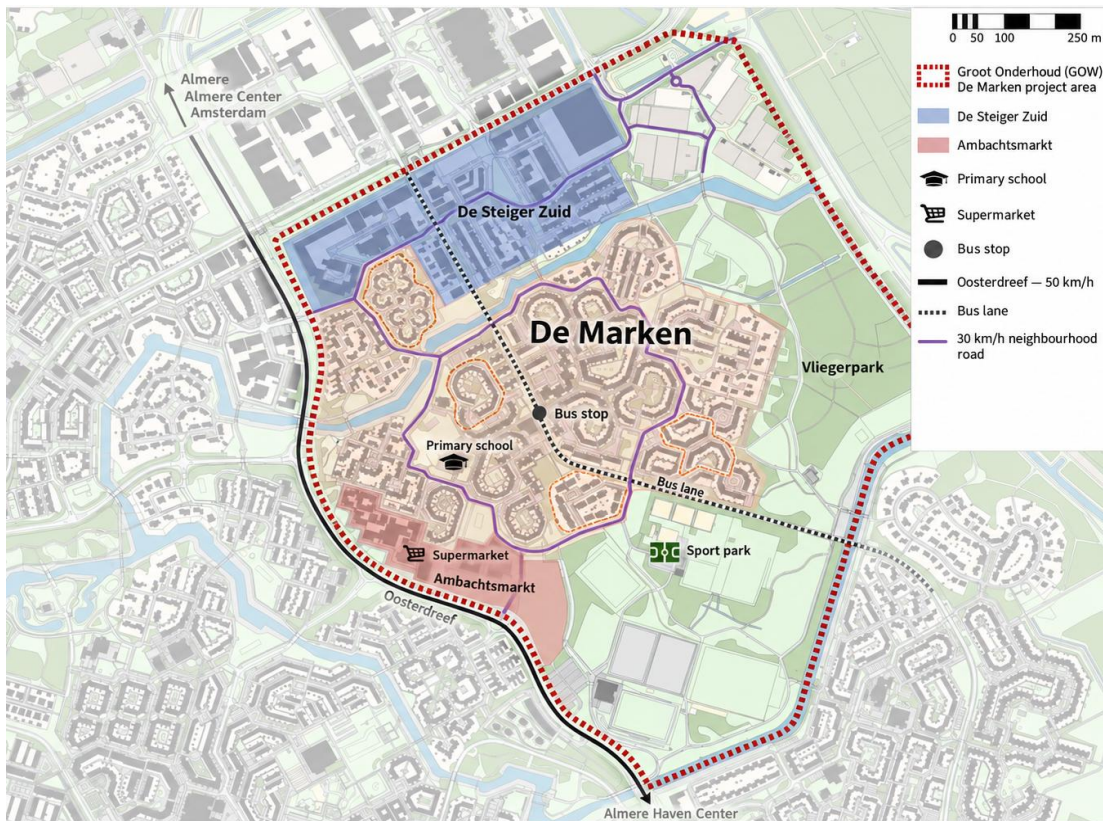
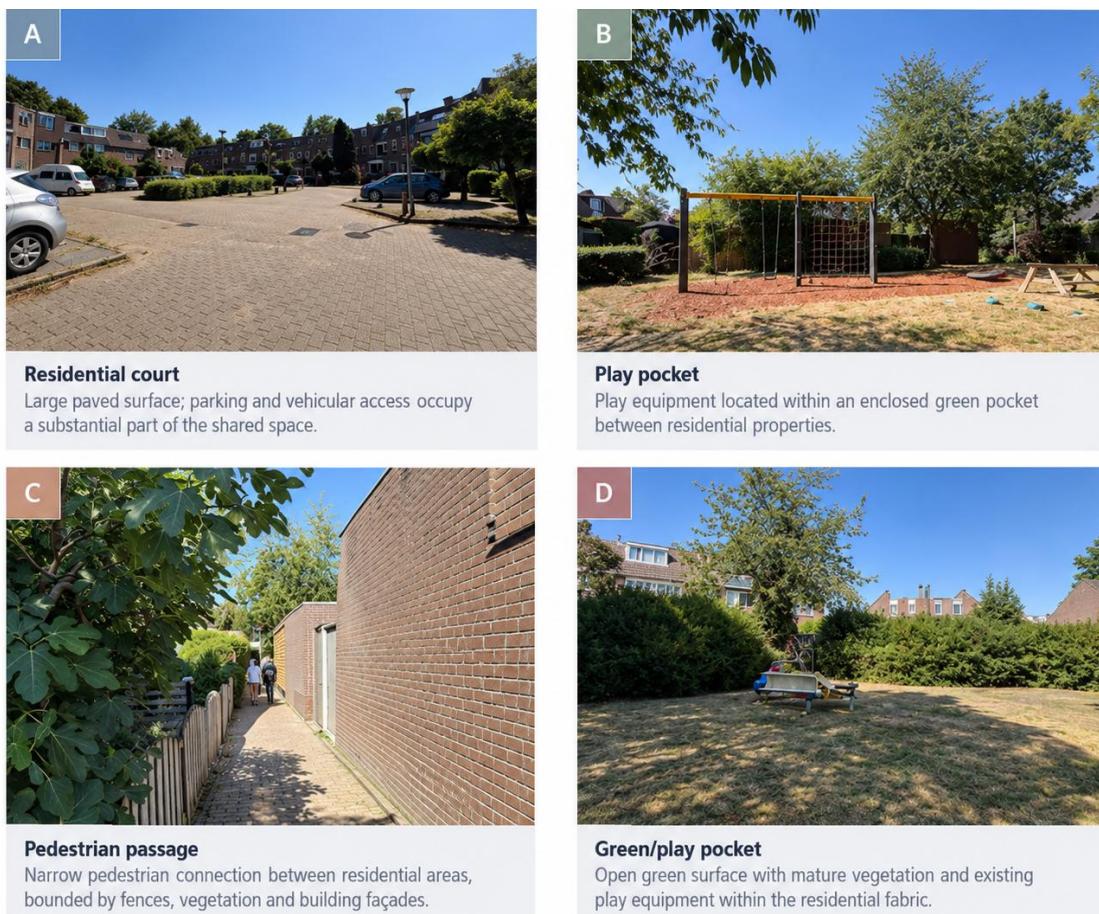


Figure 4. Selected field evidence from the study area included in the frozen evidence corpus



Source: Authors' field observations, 2026.

The resulting challenge was to identify how planned renewal could generate wider neighbourhood value while remaining grounded in De Marken's existing conditions and compatible with Almere's longer-term objectives.

3.3. Evidence base and uncertainty

The Frozen Evidence Corpus was assembled before execution and held constant across Conditions A, B and C (Table 2; Supplementary Material A2). It combined municipal programme and policy documents, neighbourhood statistics, spatial and thematic maps, field photographs and contextual research. Together, these supported early-stage integrated assessment while leaving genuine uncertainty concerning detailed feasibility, selected ownership interfaces, costs, actor commitments and pending municipal decisions.

Table 2. Frozen evidence corpus shared across experimental conditions

Evidence category	Representative source(s)	Spatial scale	Evidence role
Spatial structure and morphology	De Marken gebied (E06); De Marken stedenbouw (E07); authors' field photographs (E15) and accompanying photo-location map (E16)	neighbourhood / pocket	spatial / visual local evidence
Public-space condition and maintenance	GOW wijkanalyse De Marken (E01); authors' field photographs (E15)	neighbourhood / street	operational / visual local evidence
Demography, housing, liveability and safety	GOW wijkanalyse De Marken (E01); Data Leefbaarometer: De Marken (E04); Data De Marken (E05)	neighbourhood / buurt	statistical / qualitative local evidence
Functions, facilities and play infrastructure	De Marken functions (E08); Facilities (E09); Playground facilities (E10)	neighbourhood / surroundings / equipment location	functional / spatial local evidence
Green, ecology and climate	GOW wijkanalyse De Marken (E01); De Marken groen (E12); Green types and trees (E13)	neighbourhood / green-blue network	environmental / spatial local evidence
Mobility and connectivity	De Marken stedenbouw (E07); De Marken mobility (E11)	neighbourhood / district	network / spatial local evidence
Land ownership and land rights	Property ground (E14)	neighbourhood / parcel	ownership / spatial local evidence
Energy and circular material flows	GOW wijkanalyse De Marken (E01)	neighbourhood / municipal system	technical / operational local evidence
Planning and policy context	Ontwerp Omgevingsvisie Almere 2050 (E02); Hoofdlijnenakkoord Almere 2026–2030 (E03)	city / neighbourhood	strategic policy context
New Town typological context	Zorgen voor New Towns: Van groeikern tot aandachtsgebied (E17)	national / typological	contextual / typological evidence

Source: Authors' compilation, 2026,
based on the sources documented in Supplementary Material A2.

No additional case-specific evidence was introduced after freezing. Where the corpus could not support a conclusion, outputs were required to retain an evidence gap or unresolved decision rather than resolve it through external information or unsupported inference. Case illustrations documented existing conditions only; the authors' reference diagnosis and proposed regenerative solution were excluded from the shared inputs. Selected field photographs are shown in Figure 4.

3.4. From planning case to controlled experimental task

The case was formalised as a controlled input package comprising the Municipal Brief, Frozen Evidence Corpus and Common Output Specification. Together, these fixed the municipal assignment, permissible case-specific evidence and required decision-facing deliverable while leaving the analytical procedure to each experimental condition.

The Municipal Brief required an evidence-led reading of De Marken as an interconnected spatial–environmental–social system; interpretation of its bloemkoolwijk morphology; assessment of GOW’s potential beyond standard physical renewal; definition of actions within GOW or requiring aligned programmes and resources; and formulation of a coherent public-realm package. Findings had to be grounded in E01–E17, with feasibility constraints, evidence gaps and unresolved decisions retained.

The Common Output Specification standardised the client-facing deliverable without prescribing method: an evidence-based diagnosis, prioritised challenges and opportunities, an integrated regeneration direction, a concrete intervention portfolio, implementation and governance requirements, and explicit treatment of evidence, assumptions and gaps. It also required evidence traceability and prohibited unsupported case-specific information or invented costs, capacities, actor commitments and quantitative outcomes.

No condition received a pre-authored diagnosis, preferred regenerative direction or final spatial solution. The complete controlled package is documented in Supplementary Material A: Municipal Brief (A1), Frozen Evidence Corpus - Source Inventory and Documentation (A2), and Common Output Specification (A3).

4. Experimental Design and Execution

4.1. Comparative research design and controlled variables

The study used an exploratory controlled comparison of three execution configurations. The municipal planning question, Frozen Evidence Corpus, model and client-facing output specification were held constant across conditions. External retrieval and substantive human intervention during generation were excluded. Condition-specific differences in methodological governance, execution configuration and timing of exposure to the client-facing output specification are summarised in Table 3.

Table 3. Experimental conditions

Variable	Condition A	Condition B	Condition C
Municipal planning question	Same	Same	Same
Frozen evidence corpus	Same	Same	Same
Model and version	Same	Same	Same
Run settings	Same	Same	Same
Client-facing output specification	Same	Same	Same
Explicit SCM specification	No	Yes - shared SCM Methodological Reference	Yes - shared SCM Methodological Reference
Execution configuration	Single prompt-response	Single-model, single-pass execution of P1–P9 + final translation	P1–P9 distributed across bounded sequential roles + final renderer
Externalised intermediate handoffs	No	No	Yes - transferred unchanged
Timing of output-specification exposure	Initial	Present initially; use deferred until after	C05 only

Variable	Condition A	Condition B	Condition C
		P1–P9	
External retrieval	No	No	No
Substantive human intervention during generation	None	None	None

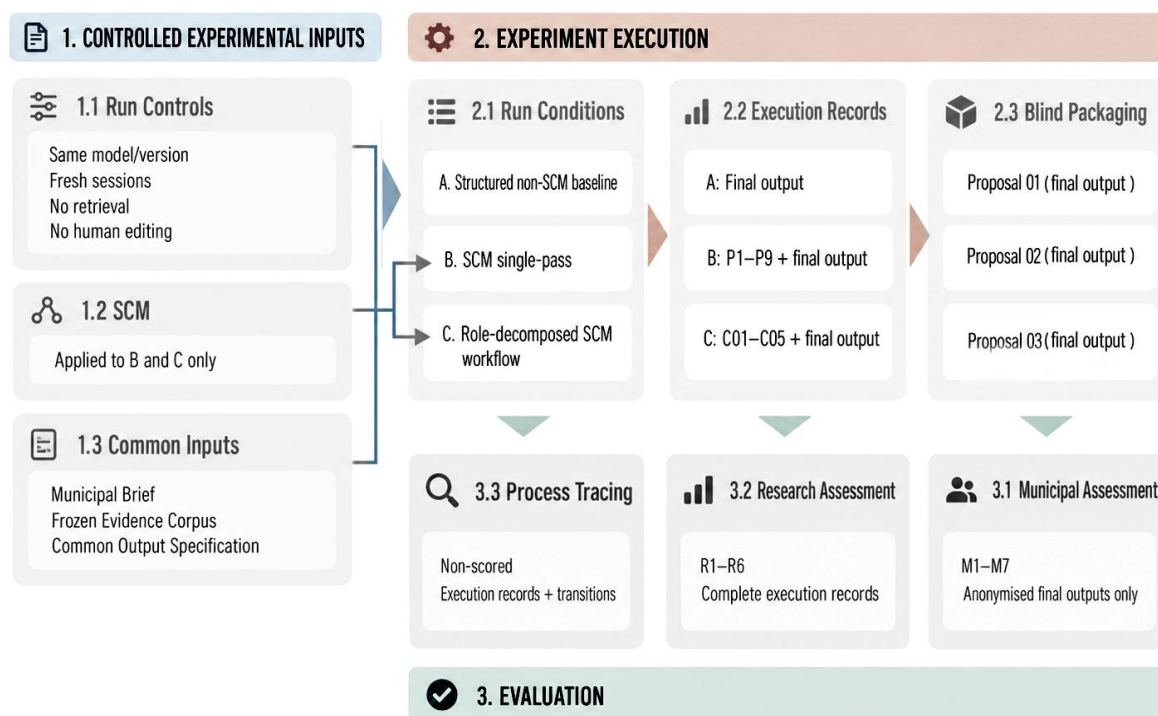
Source: Author 1, 2026

4.2. Experimental protocol and run controls

Each condition was executed once using fresh Grok 4.6 sessions (xAI chat interface, Expert mode). A and B each comprised one prompt submission and one model response; C comprised five predefined role executions (C01–C05), each in a fresh chat, with outputs transferred unchanged between successive roles.

The case inputs and condition-specific prompts were locked before execution. Web search, external connectors and code execution were disabled; decoding parameters were not user-configurable and followed the interface default. No prior conversation, project memory or previous De Marken outputs were available in context. No substantive human interpretation, correction or editing occurred during generation, and completed outputs were preserved unchanged. The complete experimental and evaluation protocol is

Figure 5. Experimental Design and Evaluation Protocol



Source: Author 1, 2026.

summarised in Figure 5.

4.3. Condition A - Structured non-SCM baseline

A received the Municipal Brief, Frozen Evidence Corpus and Common Output Specification, but no SCM Methodological Reference, SCM-specific analytical sequence or reasoning rules. The model determined its own analytical structure and produced one response in the shared client-facing format.

4.4. Condition B - SCM-encoded single-pass execution

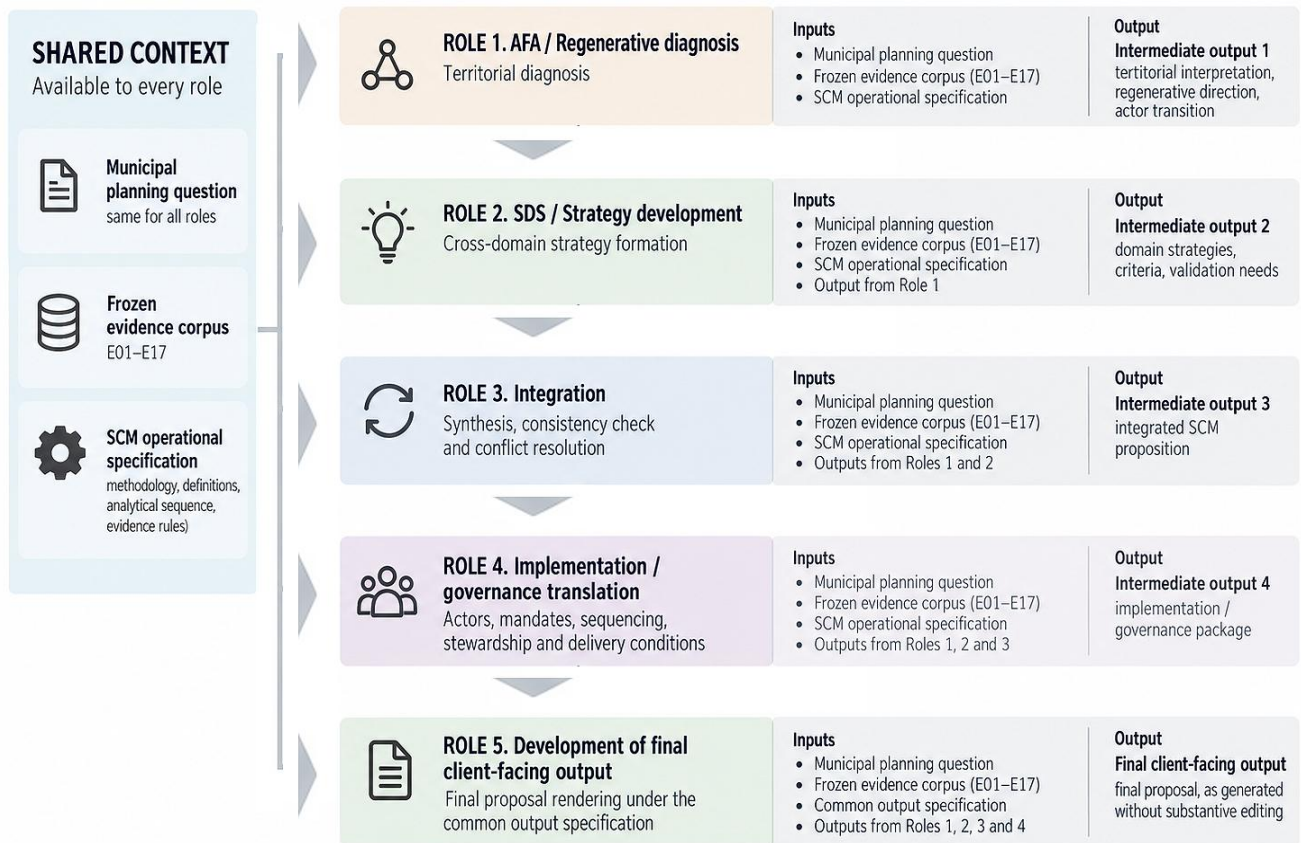
B received the same inputs plus the frozen SCM Methodological Reference. Its required products P1–P9, evidence-use and sequencing rules were encoded in one prompt. The brief, corpus, SCM specification and Common Output Specification were issued together; no second prompt, intermediate inspection or revision pass was permitted.

The single response contained two parts within the same generation: an internal SCM execution record (P1–P9) and its translation into the common client-facing format. Part II was not a separate model call.

4.5. Condition C - Role-decomposed SCM workflow

C executed the same methodology-derived sequence through five bounded sequential roles with unchanged handoffs (Figure 6; Table 4): C01 produced P1–P4; C02, P5–P7; C03, P8; C04, P9; and C05 rendered the final client-facing output. Each role was executed in a fresh

Figure 6. Condition C workflow: bounded methodological roles and sequential manual unchanged handoffs.



Source: Author 1, 2026

chat and received all preceding role outputs separately and unchanged.

Table 4. Execution of the common methodology-derived sequence in Conditions B and C

Methodological / translation sequence	Condition B - single-pass execution	Condition C - role-decomposed execution
P1–P4 - AFA territorial interpretation	Part I - continuous model context	C01 - AFA Territorial Interpreter
P5–P7 - SDS development		C02 - SDS Strategy Developer
P8 - SCM integration		C03 - SCM Integrator
P9 - Implementation & governance translation		C04 - Implementation & Governance Translator
Final client-facing output	Part II - same model response	C05 - Final Output Renderer

Source: Author 1, 2026

C01–C04 received the Municipal Brief, Frozen Evidence Corpus and SCM Methodological Reference, but not the Common Output Specification. After P1–P9 were completed, C05 received all preceding role outputs separately and unchanged, together with the Common Output Specification and the corpus for verification. C05 translated the completed SCM execution into the shared client-facing format without reopening the diagnosis, regenerating strategies or introducing new substantive analysis. C was manually orchestrated at handoffs only; no human gates were used. It therefore represents a role-decomposed SCM workflow rather than an autonomous multi-agent system.

4.6. Prototype status and TRL positioning

The tested configuration is positioned at TRL 5. It was executed in the live De Marken renewal context, on frozen municipal evidence base and a decision-facing planning task, and assessed by an urbanist of the Municipality of Almere with direct knowledge of the case.

5. Evaluation Framework

The experiment used two separate scored assessments and one non-scored process analysis: complete-record research evaluation (R1–R6), condition-blind municipal-practice evaluation of final client-facing outputs (M1–M7), and process tracing of how methodological reasoning was developed, preserved and translated.

5.1. Regenerative-reasoning assessment

Six criteria derived from Section 2 assessed territorial relationships, place-specific regenerative direction, spatial differentiation, cross-domain coordination, translation into performance and delivery, and governance capacity for longer-term transformation (Table 5). They assessed these qualities rather than SCM terminology; Condition A could therefore meet every criterion without using SCM-specific concepts.

Table 5. Research evaluation criteria for regenerative urban-planning reasoning

Code	Criterion	Evaluation focus
R1	Relational territorial intelligence	Whether the territory is understood as an interconnected socio-ecological and spatial system, including relationships among morphology, ecology, infrastructure, flows, everyday use and actors.
R2	Place-specific regenerative direction	Whether the territorial reading generates a coherent transformation direction from local relationships, conflicts, assets and latent capacities rather than a generic set of improvements.

Code	Criterion	Evaluation focus
R3	Place-responsive spatial differentiation across socio-ecological systems	Whether differences in morphology, ecology, hydrology, flows, use and local capacity produce differentiated spatial roles and interventions.
R4	Cross-domain coherence, synergies and trade-offs	Whether strategies operate as an integrated system and address both reinforcing effects and material conflicts across domains.
R5	Accountable translation from territorial direction to strategy and performance	Whether territorial direction is traceably translated into operational strategies, implementation conditions and performance or validation criteria.
R6	Regenerative governance, participation and delivery capacity	Whether the proposal develops the governance, participation and stewardship relationships required to sustain territorial transformation over time.

Source: Author 1, 2026.

The unit of analysis was the complete generated record: A's final output; B's P1–P9 record and client-facing package; and C's C01–C05 record, including P1–P9 and the final package. Each criterion was scored once per condition for the quality, consistency and traceability of relevant reasoning. Criteria were equally weighted and averaged; output length and number of artefacts were not rewarded. Author 1 scored the fixed records. This was an analytical research assessment, not independent external validation (Supplementary B1–B2).

5.2. Condition-blind municipal decision-usefulness assessment

Author 2, a Gemeente Almere practitioner with direct knowledge of De Marken and GOW, scored only the three anonymised client-facing outputs. She did not design the protocol, write prompts, run the model or know condition labels until scoring was complete (Supplementary C1–C3).

Seven practice criteria assessed municipal decision usefulness without reference to SCM constructs (Table 6). For each output, the evaluator stated its organising logic, scored M1–M7, and recorded its main strength, weakness and any unsupported claim.

Table 6. Municipal-practice evaluation criteria

Code	Criterion	Evaluation focus
M1	Problem diagnosis and prioritisation	Whether the output identifies the main challenges and opportunities and distinguishes priorities clearly.
M2	Evidence grounding and traceability	Whether claims and recommendations are supported by and traceable to the supplied evidence.
M3	Systemic integration	Whether spatial, environmental, social and implementation issues are connected through shared relationships rather than treated separately.
M4	Strategic coherence	Whether the output presents a clear overall direction in which proposed interventions derive from and reinforce one another.
M5	Context specificity and actionability	Whether the analysis is specific to De Marken and translated into relevant, implementable directions.
M6	Treatment of uncertainty and evidence gaps	Whether supported conclusions are distinguished from assumptions and important validation needs are identified.
M7	Overall decision usefulness	Whether the output provides a credible and useful basis for municipal decision-making.

Source: Authors, 2026.

After scoring, she compared the outputs on overall decision-making basis, integration and coherence, and the balance between long-term neighbourhood value and realistic GOW implementation, and recorded whether length or detail affected judgement (Supplementary B3–B4).

5.3. Process-tracing analysis of methodological preservation and translation

Process tracing complemented the two scored assessments. It was not a third scoring instrument and did not alter the research or municipal-practice scores. Prompts and model-call structure were not coded.

The unit of observation was the complete generated record: Condition B: P1–P9 plus the final output; Condition C: C01–C05 plus the final output. Condition A contained no prescribed intermediate SCM products, so tracing was limited to B and C and used to interpret the B–C contrast.

Two unscoring observations were made.

1. Methodological operations (O1–O7; Table 7). Each operation was coded as:

- present - executed as a distinct step and producing a usable analytical product;
- partial - present in the record but not organising the next step;
- absent - not identifiable in the record.

2. Evidence status. The SCM specification required B and C stage outputs to mark claims as:

- E — directly supported by the frozen corpus;
- I — reasoned synthesis grounded in that corpus;
- G — a conclusion the corpus could not warrant.

Table 7. Process-tracing operations (O1–O7)

ID	Operation	Transition observed
O1	Lock of a single Pattern after Stage I	AFA → SDS continuity
O2	Inheritance of that Pattern into SDS without silent replacement	AFA → SDS continuity
O3	Portfolio-level test of Pattern realisation versus domain self-optimisation	Transversal integration
O4	Recombination of strategies around shared mechanisms	Transversal integration
O5	Two-layer performance logic (works indicators and Pattern tests)	Transversal integration
O6	Translation into delivery classes, freeze/go conditions and stewardship rules	Implementation and governance translation
O7	Rendering into the Common Output Specification without reopening substantive analysis	Execution record → client-facing output

Source: Authors, 2026.

Tracing recorded whether a missing warrant was classified, carried forward, converted into a delivery condition, or rewritten as an ordinary claim. Human gates were not part of the protocol. Where a G item would have required a return to professional or municipal judgement, that return was noted as unexecuted.

Supplementary Material D1 reports the O1–O7 codebook and the Vrijmark probe. Supplementary Material D2 reports the G register across C01–C05. Section 6.2 reports the operation codes, evidence status and residual decisions.

6. Results

6.1. Overall comparative results

After the municipal evaluation, the allocation key was opened and the anonymised proposals were mapped to Conditions A–C. Table 8 reports both scored assessments. They evaluate different units and constructs and are not numerically commensurable: the research evaluation scores the complete execution record; the municipal evaluation scores only the final client-facing package (Supplementary B2, B4).

The two streams produced different orderings. Research means were 3.83 (A), 4.50 (B) and 5.00 (C) - $C > B > A$. Municipal means were 3.86 (A), 3.71 (B) and 4.86 (C) - $C > A > B$. The A–B relation reversed with the unit of assessment: B was stronger on complete-record reasoning, especially R1–R2; A was stronger as a final package on diagnosis and actionability (M1, M5), while both scored 3/5 on overall usefulness (M7). Remaining B–C differences sat downstream, in integration, performance translation and delivery (R4–R6; M3–M7; Section 6.2).

Table 8. Criterion-level comparative evaluation of Conditions A, B and C

Criterion	A	B	C	Comparative finding
Panel A. Research evaluation of regenerative urban-planning reasoning				
R1 Relational territorial intelligence	3	5	5	B and C both establish coherent place-specific relational models; A remains closer to an integrated thematic diagnosis.
R2 Place-specific regenerative direction	4	5	5	B and C more explicitly derive the transformation direction from place-specific territorial relationships.
R3 Place-responsive spatial differentiation	5	5	5	All three strongly differentiate intervention according to local spatial, ecological and institutional conditions.
R4 Cross-domain coherence, synergies and trade-offs	4	4	5	C most fully reorganises domain strategies into integrated packages and retains conflicts as explicit conditions or decision dependencies.
R5 Accountable translation to strategy and performance	3	4	5	Performance translation remained the most differentiating dimension: B established the chain, while C more fully consolidated it into integrated packages and explicit performance/validation tests.
R6 Regenerative governance, participation and delivery capacity	4	4	5	C most fully translates governance and stewardship into a cross-package delivery architecture with sequencing, dependencies and differentiated responsibilities.
Mean / 5	3.83	4.50	5.00	C > B > A
Panel B. Condition-blind municipal-practice evaluation				
M1 Problem diagnosis and prioritisation	5	3	4	A provided the clearest diagnosis and prioritisation; B was less explicit in signalling priorities.
M2 Evidence grounding and traceability	5	5	5	All three proposals were assessed as strongly evidence-grounded.
M3 Systemic integration	3	4	5	Systemic integration increased from A to B to C.
M4 Strategic coherence	3	4	5	C most clearly connected interventions to one coherent overall direction.
M5 Context specificity and actionability	4	3	5	C was judged most actionable in the De Marken/GOW context; B was weaker in practical translation.
M6 Treatment of uncertainty and evidence gaps	4	4	5	C most consistently retained unresolved dependencies and validation needs.

Criterion	A	B	C	Comparative finding
M7 Overall decision usefulness	3	3	5	C provided the strongest basis for municipal decision-making.
Mean / 5	3.86	3.71	4.86	C > A > B

Source: Authors, 2026.

6.2. Process tracing: what was preserved and where it was lost

Stage I (AFA) produces a territorial interpretation, Pattern and Scenario. Stage II (SDS) turns that direction into domain strategies, criteria and metrics. Integration is not a third substantive stage. It tests whether the SDS portfolio still serves the Pattern, or whether domains have begun to self-optimize. Condition A had no prescribed intermediate products, so tracing is limited to B and C.

Seven operations (O1–O7) were coded on the complete records as present, partial or absent (Table 9). Cell-level evidence and the Vrijmark probe are reported in Supplementary Material D1.

Table 9. Process operations in Conditions B and C

ID	Operation	B	C
01	Lock of a single Pattern after Stage I	Present	Present
02	Inheritance of that Pattern into SDS without silent replacement	Present	Present
03	Portfolio-level test of Pattern realisation versus domain self-optimisation	Partial	Present
04	Recombination of strategies around shared mechanisms	Partial	Present
05	Two-layer performance logic	Partial	Present
06	Delivery classes, freeze/go conditions and stewardship rules	Partial	Present
07	Rendering into the Common Output Specification without reopening substantive analysis	Partial	Present

Source: Author 1, 2026.

The codes split the two contrasts:

- **01–02 are shared.** B and C lock and carry a Regenerative Master Pattern; A does not use such a model as an organising frame.

- **03–07 distinguish B from C.** In B these operations occur inside a continuous P1–P9 sequence: synergies and conflicts are stated, but the SDS catalogue remains the organising frame, and part of that structure is compressed in the final output. In C they occur as a separate integration-and-render sequence: the portfolio is regrouped around Pattern mechanisms, delivery rules are applied, and the render pass does not reopen analysis.

The B–C contrast is therefore single-pass execution of SCM against role-decomposed staged execution of the same methodology.

One local case shows the difference that survives into the municipal brief. Both conditions treat Vrijmark as a coupled physical–social problem rather than a free-standing greening task. Only C keeps selective non-intervention as the Pattern-consistent option if works and stewardship cannot be paired, and that rule remains in the final output. In B the same tension is present in the record and thins out in the brief.

The pairing rule is an evidence-status problem as much as an integration problem. A methodology-governed record must keep missing warrants classified

- Condition C carried the G register from Pattern lock through rendering without inventing closure (Supplementary D2).

- Condition B identified uncertainty within a continuous record, but did not preserve that status through explicit staged handoffs.

On M6, C kept missing warrants classified more consistently than A or B. The classifications serve the territorial reading: a Pattern-consistent rule is not rewritten as a completed measure when the frozen corpus cannot evidence the pairing it requires. Residual decisions in the final package are therefore G items deliberately left classified. They were generated and preserved by the workflow, not validated through a human gate (Section 5.3).

7. Discussion

The run separates three questions that are often scored as one. Can a model write a plausible public-realm package? Does an explicit methodology change how the territory is read as a living system rather than as assets and deficits? Does that reading still organise intervention after it has been turned into strategies, delivery logic and a client-facing brief — including which relationships to repair, intensify, or leave low-intensity?

Those are different objects. Generation was not the scarce capacity. What had to be tested was whether a place-based relation could remain the organising rule while a conventional municipal package was produced.

7.1. *Specification and execution govern different layers*

Specifying professional reasoning does not ensure that a place-based reading survives in the municipal package. A methodology can enter an LLM workflow as knowledge supplied to the model, or as an executable specification of what may count as evidence, which transitions are required, and which commitments may not be closed. The first is context. The second is a constraint on transformation. Encoding SCM structured how the problem was reasoned about. Execution architecture determined whether that reasoning could be re-tested, translated and rendered without being silently rewritten.

The A–B contrast isolates specification. With the brief, corpus, model and output requirements held constant, SCM was associated with stronger relational territorial intelligence and a more explicit regenerative direction (R1–R2). It did not produce a different place, and it did not raise spatial differentiation: R3 was already at ceiling in A. B and C locked the same kind of Pattern and carried it into SDS (O1–O2). Splitting Stage I into further generative roles was not required for that lock.

The B–C contrast isolates execution of the same specification. Once the Pattern existed, the single pass kept integration, delivery and unresolved dependencies inside one continuous record. The staged run made them separate operations (O3–O7). B outranked A on the complete record and did not outrank A on overall decision usefulness (M7 = 3 in both). The limit of specification-without-staging was the package, not the reading.

AFA and SDS therefore need not be separate agents. They do need to exist as separate analytical products before integration. Control sits not only in inspection of the finished brief, but in which later step is allowed to change the territorial claim

7.2. *Regenerative continuity is re-validation of a living reading*

Regenerative continuity is not retention of prior context. It is a repeated question: as the proposition changes form, do the same relationships still organise the intervention (Reed, 2007; Mang and Reed, 2012)? Metrics do not define that direction. They test whether strategies derived from the Pattern still serve the relations from which the Pattern was read.

In De Marken those relations were read as two interconnected networks: a slow continuum of nested hofjes, water and park as the primary structure, and faster seams as permeable connections to it. Public-realm renewal that replaces a surface can still impoverish that system if the room, the circuit and the seam are treated as separate works. B and C established the same kind of Pattern (R1–R3; O1–O2). The difference appeared once

domain strategies existed. C returned the SDS portfolio to the Pattern and asked whether sectoral measures still realised the territorial proposition or had begun to self-optimize (O3–O5; R4).

That test changes what counts as success. Accessibility can outrank reuse of uneven stone on the walking circuit. Water retention cannot convert a hof into a fenced basin. Physical upgrade at Vrijmark can be withheld where works and stewardship cannot be paired. Both conditions treated Vrijmark as a coupled physical–social problem, not a greening task. Only C kept selective non-intervention as the Pattern-consistent option, and that rule remains in the brief (Supplementary D1). Withholding is not a shortfall of ambition. In a living public-realm system it can be the act that protects the relation the project claims to regenerate.

The run does not show that these tests are correct without field validation. It shows that they can be represented, asked again, and preserved — which is the condition under which a municipal officer can join the reasoning rather than receive a completed picture of the future (Cugurullo and Xu, 2025).

7.3. What counted as success

The usual urban-AI metrics — speed, fluency, completeness of prose, model autonomy — measure the wrong object. Two scored assessments already diverge, and they should. The research evaluation asked whether territorial relations still organised reasoning in the execution record. The municipal evaluation asked whether the final artefact could support a GOW decision without that record.

The orderings differ (Table 8). Integrity of the record: $C > B > A$. Means of the brief: $C > A > B$. Overall decision usefulness is M7, not the municipal mean: A and B are equal and low; only C reaches 5. **The A–B reversal across the two assessments is itself a result.** A more coherent methodological record is not automatically a more usable brief. A clearer diagnostic package is not automatically a regenerative reading that survives translation.

Success in this pilot is therefore fourfold and conjunctive: the reading became territorial; the Pattern still organised the package; gaps remained classified; and the package could be used, amended or refused without reconstructing an opaque chain. Usefulness without the Pattern tests would be the success of A. A high research score that dies in the brief would be the success of B. Neither is sufficient.

7.4. The dossier is interrogated, not completed

The municipal dossier can warrant assets, deficits and delivery packages. A regenerative reading asks for relations the file does not already hold: pairing of works and stewardship, ownership of hofjes, aftercare, and the scope of GOW against the fast seams. Methodology-governed reasoning does not force the project to fit the available file. It uses the file to test what each step can warrant and leaves visible what must still be obtained.

Missing warrant is not an information failure to be hidden by a complete-looking output. It is actionable state. Condition A, with no prescribed status rule, can rewrite a missing warrant as an ordinary proposal. B named uncertainty inside a continuous record and did not carry that status through staged handoffs. C marked gaps at Pattern lock (C01), wrote dependent strategies as qualified (C02), converted part of the register into portfolio rules — pairing and selective non-intervention (C03), placed the remainder on freeze/go and actor memos (C04), and left the register visible in the brief without invented closure (C05). No human entered the loop. The gates this run can locate — pairing, ownership, aftercare — were not executed.

The B–C difference does not show that regenerative planning needs more specialised agents. B and C were already equivalent on R1–R3. Differences sat where an established proposition changed epistemic or institutional form (O3–O7; R4–R6; M7). Integration may re-test the portfolio against the Pattern. Delivery may write sequencing, stewardship and freeze/go. Rendering may reorganise the proposition for GOW review; in this protocol it

may not reopen diagnosis or close a G item. That ban on reopening is an experimental constraint, not an SCM principle. The next cycle should return those G items to judgement and may reopen Stage I.

8. Limitations

The comparison is bounded by design: one neighbourhood, one model family, one fresh run per condition. That bound is what made specification and execution separable from case variation. Stochastic variation and cross-model robustness were not assessed.

Author 2 scored the anonymised packages with direct knowledge of De Marken and GOW. The scoring was condition-blind and role-separated, not institutionally independent. Author 1 designed the protocol, ran the model and scored the research records. Condition C was manually orchestrated. No human gate was executed.

The study evaluates generated decision-support artefacts, not municipal adoption or built outcomes. Findings concern whether a place-based regenerative reading can be structured, preserved and translated under the controls reported here. Condition C used sequential isolated model calls; Condition B used one continuous call. That difference belongs to the execution configuration as tested.

9. Conclusions

A regenerative reading of place survives a municipal package only if the relationships through which the territory was interpreted still organise intervention — including which actions to intensify, repair or withhold.

In this experiment, encoding SCM as an executable specification was associated with that reading in the execution record (A→B; R1–R2). Executing the same specification through a staged workflow was associated with keeping the reading available in the decision-facing artefact (B→C; R4–R6; M3–M7). Those are different layers. Research scores ranked C > B > A. Municipal scores ranked C > A > B. Only C was judged decision-ready (M7 = 5). Four claims follow.

1. **Specifying professional reasoning does not ensure its preservation through transformation.** SCM structured territorial interpretation and regenerative direction. Staging became consequential once that reasoning had to be integrated, delivered and rendered.
2. **Methodological continuity depends on re-testing relationships,** not on retaining prior context. B and C locked the same kind of Pattern. Only C made portfolio-level re-testing a distinct operation and kept a Pattern-consistent refusal rule visible in the brief.
3. **A municipal package is produced through delivery and rendering, not through generation alone.** Delivery may write sequencing, freeze/go and stewardship. Rendering may change representation. In this protocol it may not reopen diagnosis or close what the dossier cannot warrant. That ban is an experimental constraint; the next cycle may reopen Stage I at the G items already marked.
4. **Methodological integrity and decision usefulness can diverge.** B outranked A on the record; A was at least as usable as B as a brief. The record and the brief should be assessed apart.

What can travel is that distinction: specification; a place-derived Pattern against which later representations are re-tested; integrator, delivery and renderer as later functions; two separate assessments. What should not travel is a multi-agent graph, an autonomous plan, or the De Marken package as a model answer. The next loop returns pairing, ownership and aftercare to judgement.

10. Supplementary Materials.

Supplementary files accompany this article and are excluded from the main-text word count.

Supplementary Material A - Controlled Experimental Case Inputs

- A1. Municipal Brief.
- A2. Frozen Evidence Corpus - Source Inventory and Documentation (E01–E17), including source notes for E05, E10, E15 and E16.
- A3. Common Output Specification.

Supplementary Material B - Evaluation Materials

- B1. Research Evaluation Rubric (R1–R6).
- B2. Completed Research Evaluation.
- B3. Municipal-Practice Evaluation Instrument (M1–M7).
- B4. Completed Condition-Blind Municipal-Practice Evaluation, including comparative judgement and length/detail check.

Supplementary Material C - Experimental Outputs

- C1. Condition A - Unedited Client-Facing Output.
- C2. Condition B - Unedited Client-Facing Output.
- C3. Condition C - Unedited Client-Facing Output.

Supplementary Material D — Process Tracing

- D1. 01–07 codebook, cell-level codes and Vrijmark probe.
- D2. Evidence-status register (E/I/G) across C01–C05.

Supplementary data and reproducibility statement. The public supplementary package provides the controlled case inputs, evaluation instruments, completed evaluations and unedited client-facing experimental outputs. The complete frozen SCM Methodological Reference, condition-specific prompt implementations and full execution records contain proprietary methodological know-how used in professional practice and are therefore not publicly released. These materials are retained by the authors in frozen, versioned form.

To support transparency and auditability while protecting proprietary methodology, the article reports the methodological architecture, P1–P9 analytical products, execution configurations, handoff rules, evidence constraints and evaluation protocol required to assess the experimental design and interpret the reported comparisons.

11. Author Contributions

Author 1: Conceptualization; SCM methodology; experimental design; development of the Municipal Brief and Common Output Specification; construction of the controlled experimental protocol; experimental execution; regenerative-reasoning assessment; process-tracing analysis; interpretation of findings; writing-original draft; writing-review and editing; project administration.

Author 2: Municipal case-study expertise; provision of municipal data and documentation forming the evidence base for the Frozen Evidence Corpus; contribution to the case-study section; review and validation of the Municipal Brief, Frozen Evidence Corpus and Common Output Specification from the municipal-practice perspective; condition-blind municipal-practice evaluation; writing-review and editing.

References

- Amershi, S., Weld, D., Vorvoreanu, M., Fournery, A., Nushi, B., Collisson, P., Suh, J., Iqbal, S., Bennett, P. N., Inkpen, K., Teevan, J., Kikin-Gil, R., & Horvitz, E. (2019). Guidelines for human-AI interaction. *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*, 1–13. <https://doi.org/10.1145/3290605.3300233>
- Bijlsma, L., Bergenhenegouwen, G., Schluchter, S., & Zaaijer, L. (2008). *Transformatie van woonwijken: Met behoud van stedenbouwkundige identiteit*. NAI Uitgevers & Ruimtelijk Planbureau.

- Cole, R. J. (2023). Transition to a regenerative future: a question of time. *Buildings & Cities*, 4(1), 457–474. <https://doi.org/10.5334/bc.333>
- Cole, R. J., Oliver, A., & Robinson, J. (2013). Regenerative design, socio-ecological systems and co-evolution. *Building Research & Information*, 41(2), 237–247. <https://doi.org/10.1080/09613218.2013.747130>
- Cugurullo, F., & Xu, Y. (2024). When AIs become oracles: generative artificial intelligence, anticipatory urban governance, and the future of cities. *Policy and Society*. <https://doi.org/10.1093/polsoc/puae025>
- de Meere, F., Verwaijen, G., Uytterlinde, M., van der Velden, J., van der Hoeven, A., Hollanders, K., van Eikeren, P., & Louter, P. (2024). Zorgen voor New Towns: Van groeikern tot aandachtsgebied. WijkWijzer.
- du Plessis, C. (2012). Towards a regenerative paradigm for the built environment. *Building Research & Information*, 40(1), 7–22. <https://doi.org/10.1080/09613218.2012.628548>
- Gemeente Almere. (2026). Globale analyse uitdagingen en kansen De Marken [Unpublished municipal working document].
- Grand, G., Tenenbaum, J. B., Mansinghka, V. K., Lew, A. K., & Andreas, J. (2025). Self-steering language models. Second Conference on Language Modelling (COLM 2025). <https://doi.org/10.48550/arXiv.2504.07081>
- Huang, Y., Li, H., Lu, Y., Wu, R., Chen, S., & Liu, C. (2026). Evidence-grounded multi-agent planning support for urban carbon governance via retrieval-augmented generation. arXiv. <https://doi.org/10.48550/arXiv.2601.11587>
- Hong, S., Zhuge, M., Chen, J., Zheng, X., Cheng, Y., Zhang, C., Wang, J., Wang, Z., Yau, S. K. S., Lin, Z., Zhou, L., Ran, C., Xiao, L., Wu, C., & Schmidhuber, J. (2024). MetaGPT: Meta programming for a multi-agent collaborative framework. International Conference on Learning Representations (ICLR 2024). <https://doi.org/10.48550/arXiv.2308.00352>
- Li, B. Z., Tamkin, A., Goodman, N. D., & Andreas, J. (2023). Eliciting human preferences with language models [Preprint]. arXiv. <https://doi.org/10.48550/arXiv.2310.11589>
- Mang, P., & Reed, B. (2012). Designing from place: A regenerative framework and methodology. *Building Research & Information*, 40(1), 23–38. <https://doi.org/10.1080/09613218.2012.621341>
- Matashova, M. (2022). Sustainable Community Masterplan: From mitigation to regeneration. In *The Palgrave Encyclopedia of Urban and Regional Futures*. Palgrave Macmillan. https://doi.org/10.1007/978-3-030-51812-7_332-1
- Matashova, M. (2021). Sustainable Community Masterplan: Integrative design process. *Architecture and Construction of Russia*, 2021(3), 34–41.
- Pan, F., Huang, X., Bi, Y., Gao, Y., Ye, Y., & Wang, H. (2025). From tools to partners: How large language models are transforming urban planning. *AI Open*, 6, 276–298. <https://doi.org/10.1016/j.aiopen.2025.11.001>
- Reed, B. (2007). Shifting from ‘sustainability’ to regeneration. *Building Research & Information*, 35(6), 674–680. <https://doi.org/10.1080/09613210701475753>
- Rijksdienst voor het Cultureel Erfgoed. (2020). Opkomst en ontwikkeling van de bloemkoolwijk.
- Shneiderman, B. (2020). Bridging the gap between ethics and practice: Guidelines for reliable, safe, and trustworthy human-centred AI systems. *ACM Transactions on Interactive Intelligent Systems*, 10(4), Article 26. <https://doi.org/10.1145/3419764>
- Solà-Morales, M. de. (1981). La identidad de la descripción. *Quaderns d’Arquitectura i Urbanisme*, (146), 12.
- Son, T. H., Weedon, Z., Yigitcanlar, T., Sanchez, T., Corchado, J. M., & Mehmood, R. (2023). Algorithmic urban planning for smart and sustainable development: Systematic review of the literature. *Sustainable Cities and Society*, 94, Article 104562. <https://doi.org/10.1016/j.scs.2023.104562>
- Zhu, H., Chen, G., & Zhang, W. (2025). PlanGPT: Enhancing urban planning with a tailored agent framework. In *Proceedings of the 63rd Annual Meeting of the Association for Computational Linguistics (Volume 6: Industry Track)* (pp. 764–783). Association for Computational Linguistics. <https://doi.org/10.18653/v1/2025.acl-industry.54>
- Zheng, Y., Xu, F., Lin, Y., Santi, P., Ratti, C., Wang, Q. R., & Li, Y. (2025). Urban planning in the era of large language models. *Nature Computational Science*, 5, 727–736. <https://doi.org/10.1038/s43588-025-00846-1>
- Fu, X., Li, C., Quan, S. J., Yigitcanlar, T., & Wasserman, D. (2025). Large language models in urban planning. *Nature Cities*, 2, 585–592. <https://doi.org/10.1038/s44284-025-00261-7>

- Wang, Y., Fu, Y., & Lyles, W. (2026). Generative AI for spatial regeneration planning: Integrating urban planning theories and ethics. *Computers, Environment and Urban Systems*, 124, 102380. <https://doi.org/10.1016/j.compenvurbsys.2025.102380>
- Wang, L., Ma, C., Feng, X., Zhang, Z., Yang, H., Zhang, J., Chen, Z., Tang, J., Chen, X., Lin, Y., Zhao, W. X., Wei, Z., & Wen, J. (2024). A survey on large language model based autonomous agents. *Frontiers of Computer Science*, 18(6), Article 186345. <https://doi.org/10.1007/s11704-024-40231-1>
- Wong, L., Mao, J., Sharma, P., Siegel, Z. S., Feng, J., Korneev, N., Tenenbaum, J. B., & Andreas, J. (2024). Learning adaptive planning representations with natural language guidance. *International Conference on Learning Representations (ICLR 2024)*. <https://openreview.net/forum?id=9eAls9L3qV>