FCTE / AUDIT & EVIDENCE REFERENCE

Technical methodology archive

The earlier technical explanation, retained as a full reference document.

Source and scope

This document reformats the complete published source. Historical results and current-tense wording belong to that source context. It is not a new rerun or independent certification. People, histories and model counts are synthetic.

Read before use

Connected synthetic histories

Persistent identities, events and relationships support the modeled city. Sources, assumptions and temporal boundaries must be retained.

Proposed controls and limits

Registry, configuration contracts and remediation controls are described in their published scope. Complete lineage, independent validation and whole-city regeneration are not established by this reference.

Request the full supporting notes from Fidelra

Full published source

Methodology · Fidelra City Twin Engine

FCTE models persistent synthetic people, households, organizations and events. Their histories connect across domains, creating a city that can be examined through its relationships, chronology and published evidence.

Read the overview

Review published checks

i

This page explains the model and its governance direction. Release evidence establishes what was tested, measured and disclosed for a particular city, domain and version.

Earlier companion overview · PDF

The existing PDF is an earlier overview; this page contains the expanded methodology.

Use & governance

The overview

The method at a glance.

• Model the city Persistent identities connect people, households, organizations and events.

• Connect its histories Shared relationships and chronology constrain how states change.

• Review each release Published checks establish what was tested. Proposed governance is identified separately.

Explore the chapters

Open a chapter for the full method, diagrams and evidence boundaries.

01 · A persistent synthetic city

The same people. A changing city.

FCTE - the Fidelra City Twin Engine - constructs a persistent synthetic city state. Residents, households, businesses and infrastructure have identities and relationships that can continue as new events occur. A record belongs to that wider model.

A synthetic resident may be linked to a household, an education history, an employment episode and healthcare activity. A business can connect work, income and geography. These relationships make it possible to ask how a state arose and what happened afterward.

Persistence creates obligations: references must resolve, transitions must be possible, and a change in one domain must remain compatible with related records. Those obligations are part of the method, while the published checks show what has actually been examined for a release.

One citizen. Many domains. Across time.

Public samples and products expose selected parts of the city. Available histories, relationships and evidence vary by city, domain and release.

02 · Evidence to configuration to generation

Make the path from assumptions to outcomes visible.

Historical records, official statistics, research and domain knowledge inform the model. Where evidence is incomplete, a model still needs explicit assumptions. Reviewers need to know which kind of input is being used, where it applies and why it was chosen.

Generation logic - the model’s writers - turns those inputs into entities, events and transitions. Configuration should express the governed assumptions and constraints; the writer still contains the logic needed to construct a connected city.

Conceptual architecture

From source evidence to a released product

Proposed governance layer. Stages marked “Governance design” describe the target registry, configuration contracts and unified change ledger.

- 01 Source evidence References, historical context and documented domain knowledge.

• 02 Governance design Parameter & evidence registry A proposed record of each material parameter's basis, scope and review history.

• 03 Governance design City × domain configuration Proposed contracts, versioned independently of writers, defining assumptions and constraints.

• 04 Generative writers Logic that constructs synthetic entities, states and events.

• 05 Persistent synthetic city Connected identities, relationships and histories across domains.

• 06 Governance design Remediation & change ledger Corrections tied to their affected scope, with a proposed unified record of changes.

• 07 Certification Defined checks, acceptance criteria, results and disclosed limitations.

• 08 STRIPED DONKEY products Selected city data, samples and release records for research and development.

Read the published release artifacts to see which controls, records and results are available for a particular product.

Follow the published evidence to assess the current release. The architecture diagram explains responsibilities; it does not report a certification outcome.

03 · Parameter & evidence governance

Explain why a parameter has its value.

A material probability, transition rule, distribution or historical boundary should have an identifiable basis. The proposed Parameter & Evidence Registry would connect that basis to the relevant city, domain and period, with its source, uncertainty and review history.

These proposed classifications describe how a parameter was established. They do not, by themselves, measure its accuracy or suitability for a particular use.

Evidence class	What it means
OBSERVED	A value taken directly from a documented measurement, survey, administrative record or official statistic.
LITERATURE	A value or relationship drawn from published research, with its source and study context recorded.
DERIVED	A value calculated from documented inputs using an explicit formula or transformation.

Evidence class	What it means
INFERRED	An estimate made from indirect or incomplete evidence, with the reasoning and uncertainty stated.
EXPERT PRIOR	An initial estimate informed by documented domain expertise, used where direct evidence is limited.
SYNTHETIC ASSUMPTION	A deliberate modelling choice used to construct or explore a synthetic scenario, rather than a measured fact.

The proposed record would retain the source date, geography, population, period of validity and confidence assessment, alongside the relevant writer version and certification gate. An official statistic from one population or period may still be a poor fit for another.

Independent configuration versions would make a changed assumption visible even when the writer’s logic stays the same. This is a governance direction, not a claim that complete record-to-evidence lineage is already published for every product.

The published Nairobi assumptions register records a selected set of assumptions. It is a useful review starting point; the proposed registry is broader in scope.

04 · Longitudinal generation

Generate transitions, not just snapshots.

A present-day state is only one point in a history. Longitudinal generation models how people, households, organizations and infrastructure change over time, retaining the events and relationships needed to interpret later states.

Education can precede an employment episode. Employment can change or end. Household membership, service access and healthcare activity can evolve alongside those changes. The chronology matters as much as the individual fields.

- 01 Education episode Learning occurs over an interval.
- 02 Qualification available Its completion belongs to a point in that history.
- 03 Employment episode Only information available at that time can inform it.
- 04 Later change A transition updates the state without erasing the earlier episode.

Conceptual sequence illustrating temporal dependencies, not a prescribed path for every resident.

Checks should examine event order, age at the event, valid episode boundaries and future-information leakage. For time-based machine-learning evaluation, features must be restricted to information available at the prediction point.

Explore a published resident journey

05 · Cross-domain coherence

A shared identity creates shared constraints.

The same synthetic entity can participate in several domains. Its household, qualifications, employment, income, mobility and healthcare history cannot be assessed as unrelated tables when the records describe the same person and period.

One synthetic resident

Identity continues through time

- Household & geography
- Education & qualifications
- Employment & income
- Telecom & mobile wallet
- Transport & mobility
- Health & insurance

Conceptual relationships across a city model. Published coverage varies by city, domain and product.

A qualification used in an employment record needs to exist at that time. A service event needs a compatible service history. A location-dependent event needs a coherent relationship to the relevant geography. A valid identifier alone does not establish those conditions.

Cross-domain review therefore examines shared references, overlapping time intervals, state compatibility and geographic consistency. A logically connected history can still simplify reality; coherence and empirical calibration remain separate questions.

Inspect a connected sample

06 · Historical epochs

Every event belongs to its time.

Institutions, infrastructure and technologies become available at different times. City-specific historical boundaries constrain which synthetic events can occur and which services or behaviours can be represented in a period.

Historical context should carry its own evidence and applicable dates. A later institutional environment should not silently be applied to earlier residents, and a broad contemporary average should not stand in for an entire historical sequence.

Recorded source values
Availability A service or technology should not appear before the relevant introduction boundary.
Eligibility An event should respect the resident's age, preceding state and applicable lifecycle rules.
Change over time Assumptions should be reviewed for the period they govern, rather than applied indefinitely.

A long historical horizon is a modelling scope. It does not establish equal evidence quality or complete real-world accuracy for every year. Historical anchors and remaining limitations should be read with the relevant city evidence.

07 · Writer governance & remediation

Fix the generating rule. Review the affected history.

A systemic defect can affect many records or related domains. Removing an inconvenient row from a final product does not address the generating logic or the history that depends on it.

The correction process is to identify the rule and affected scope, correct the writer or its assumptions, remediate the affected historical state, and add a lasting check against recurrence. Certification must be rerun after a material mutation.

- 01 Identify the defect and affected scope
- 02 Correct the writer or assumption
- 03 Remediate affected history and links
- 04 Add checks and rerun certification

• 05 Record the change and release evidence

↻ Findings feed the next review.

Correction lifecycle. A published change record and its supporting results establish which work has been completed.

A formal, unified change ledger is part of the governance design. It would connect the defect, affected period, writer change, repaired state and subsequent checks. Existing release notes and audit findings remain the place to inspect reported changes and open issues.

This process does not imply that every historical defect has been found or repaired. A completed correction claim needs a defined scope and evidence of the resulting checks.

For a scoped example, read the Nairobi health temporal cascade audit alongside its affected records, checks and findings.

Nairobi release notes

Lagos release notes

08 · Certification framework

State what a check establishes.

Certification is a framework of defined questions, acceptance criteria and recorded outcomes. It becomes useful when a reader can identify the tested release, see the scope and distinguish a result from a broader claim about realism or fitness for use.

Three questions that need separate answers

• 01 Configuration conformance Did the writer generate what the specification asked for? Matching an intended distribution demonstrates conformance to that specification.

• 02 Empirical calibration Is the configured value or range defensible against external evidence for this population, geography and period? A matching output alone cannot answer this.

• 03 System coherence Do the generated entities, events and relationships remain compatible across the city, including their chronology and cross-domain dependencies?

Seven areas of review

The framework below describes review areas, not a declaration that every area is implemented or passed for every product. ML utility and privacy evaluation require their own published evidence.

Recorded source values
Structural Required entities, valid references, identity continuity and population or state consistency.
Statistical Distributions and aggregate behaviour against stated expectations, with external comparisons identified separately.
Longitudinal Lifecycle order, episode boundaries, historical constraints and future-information leakage.
Cross-domain Compatible relationships and states where domains refer to the same entities and periods.
ML utility Performance on a defined task, with the evaluation design, comparison baseline and limitations available for review.
Privacy Assessment of relevant disclosure or identification risks for a defined product and use context.
Provenance Traceable sources, assumptions, generation context, changes and release identity.

A PASS means a defined check met its stated acceptance criteria for the identified release. It does not establish that the data is reality, suitable for every task, independently reviewed or free of privacy risk.

Keep the level of review explicit

• Engine checks: Automated tests executed by FCTE against a release; review the published scope, stated criteria and reported outcomes.

• External benchmark comparisons: selected synthetic measures compared with stated reference families, including differences and limitations.

• Independent review: only claimed when an external institution has reviewed a defined scope; FCTE certification alone does not establish this.

• Customer evaluation: suitability assessed against the customer’s intended use and empirical context.

Missing evidence, warnings and failed checks need to remain visible. Read the measured results and limitations alongside the status, including where a review area has not been evaluated.

Review measured checks and open findings

09 · Reproducibility & provenance

Keep the history of the product with the data.

A reproducible generation claim needs recorded inputs and a defined run context. Relevant evidence includes the city and domain scope, source and assumption versions, writer version, configuration version, random seed where applicable, generation run and subsequent changes.

The proposed governance layer would connect those artifacts more explicitly. A seed alone does not establish reproducibility if the implementation, dependencies or inputs have changed. Exact regeneration should only be claimed where the necessary artifacts and controls are available.

- Record
- Generator
- Configuration
- Evidence
- Certification

Target provenance path: the ability to trace a material synthetic result to the logic, assumptions, evidence and checks behind it.

Release identity is a related but distinct question. A manifest and verification value can establish whether a covered file matches a recorded artifact. Inspect signature state separately. A matching file does not prove empirical accuracy or certification across all seven review areas.

For published material, keep the city, domain, version, release manifest, verification result, data dictionary, assumptions and disclosed limitations with any downstream analysis. Review a refreshed dataset as a new version and assess the effect of reported changes.

The developer quick start demonstrates published sample replay. Whole-city regeneration, empirical fit and model performance require their own evidence.

Verify a release

Reproduce published sample checks

10 · What synthetic does - and does not - mean

A model for experimentation. A boundary around its claims.

STRIPED DONKEY records describe generated people, households, organizations and events. A synthetic resident is not a particular real individual, and a generated record cannot establish a real person’s identity, circumstances or behaviour.

Synthetic population counts are model counts, not official census estimates or direct measurements of a real city. A scenario outcome is conditional on its inputs and mechanisms; it should be kept distinct from a forecast or an observed result.

Connected synthetic environments can support software testing, research, teaching and model development. Utility for a specific task still needs evaluation. Reduced reliance on identifiable personal records does not remove the possibility of biased assumptions, misuse, misinterpretation or other risks.

The public review boundary includes data models, representative relationships, samples, outcomes and audit findings. Internal construction methods remain controlled. The evidence available for a product should support its stated claims without requiring a promise of perfect realism.

Use the product’s licence, scope, evidence and limitations to assess an intended application. Real-world evidence remains essential for challenging assumptions and deciding whether a synthetic result is useful.

Read use, governance and boundaries

Keep the supporting record

Ask Fidelra for the report context, applicable release, source references and full supporting notes. Availability and rights for a paid data package are confirmed in the written quote.

Request full notes and supporting artifacts

fidelra.africa@fidelra.ai

stripeddonkey.com

Original HTML source is attached for reference. Public data-quality datasets are also attached where applicable.

Source page: https://stripeddonkey.com/archives/method-evidence

Captured source SHA-256: 76ced428e4ab79216731c16047b84b634ed1fb60dea5de981991270169c42fb8