



TECHNICAL WHITEPAPER

Kinetiq Engine

Hybrid Generative Motion Synthesis
via Ecological Interaction Data

Isaías Reis Verdin

Thirdrez Labs

Espírito Santo, Brazil

contact@thirdrez.com

Document Number:	TRZ-2025-WQ
Version:	1.2
Status:	Production Ready
Date:	December 2025

Contents

List of Abbreviations	5
1 Introduction	6
1.1 Abstract	6
1.2 The Gelatin Problem	6
1.2.1 Origin Story: The Production Gap	6
1.2.2 The State of Generative Motion (December 2025)	7
1.2.3 Contributions	7
2 Ecological Interaction Data	8
2.1 The Problem with Laboratory Mocap	8
2.2 The Ecological Alternative	8
2.3 Ethical Considerations	9
3 System Architecture	10
3.1 High-Level Pipeline	10
3.2 Stage 1: AI Synthesis	11
3.3 Stage 2: WebAssembly Conversion	11
3.4 Stage 3: Physics Refinement	12
3.4.1 Foot Contact Detection and Locking	12
3.4.2 Joint Limit Enforcement	12
3.4.3 Smoothness Optimization	12
3.5 Stage 4: Production Export	12
4 Style LoRAs	13
4.1 The LoRA Approach	13
4.2 Available Style Modules	13
4.3 Gender Expression	13
5 Privacy-Preserving Architecture	14
5.1 The Local-First Principle	14
5.2 Privacy Guarantees	14
6 Kinetiq Editor	15
6.1 Feature Overview	15
6.2 Roadmap Features	15
6.3 Platform Targets	15
7 Benchmarks and Evaluation	16

7.1	Artifact Reduction	16
7.2	Performance Characteristics	16
8	The Human-in-the-Loop Philosophy	17
8.1	AI as Collaborator, Not Replacement	17
8.2	The Partnership Model	18
9	Conclusion	19
A	Technical Specifications	20
A.1	Supported Skeleton Standards	20
A.2	WebAssembly Module Specifications	20
	References	21

List of Figures

3.1	Kinetiq Engine System Architecture: From text prompt to production-ready animation asset. The pipeline shows server-side AI synthesis followed by client-side WASM refinement and export.	11
8.1	The Kinetiq Partnership Model: AI generates infinite variation from text prompts; humans apply artistic judgment and refinement to produce production-ready assets.	18

List of Tables

1.1	Common Artifacts in Generative Motion (The Gelatin Problem)	7
2.1	Laboratory Mocap vs. Ecological Data	8
3.1	Supported Export Formats	12
6.1	Kinetiq Editor Development Roadmap	15
7.1	Artifact Reduction Metrics	16
7.2	Client-Side Processing Latency	16
A.1	Supported Skeleton Standards	20
A.2	WASM Module Characteristics	20

List of Abbreviations

AI	Artificial Intelligence
BVH	Biovision Hierarchy (motion capture format)
DCC	Digital Content Creation
FBX	Filmbox (Autodesk animation format)
FPS	Frames Per Second
GLB	GL Transmission Format Binary
GPU	Graphics Processing Unit
IK	Inverse Kinematics
LoRA	Low-Rank Adaptation
MDM	Motion Diffusion Model
PVE	Persistent Virtual Environment
SMPL	Skinned Multi-Person Linear Model
WASM	WebAssembly

Chapter 1

Introduction

1.1 Abstract

While generative motion models can synthesize human movement from text, they consistently underperform in production environments. Outputs frequently suffer from critical physics failures: foot sliding, joint hyperextension, and mesh interpenetration. We categorize these failures as the “**Gelatin Problem**”—motion that lacks weight and deterministic contact.

This paper introduces **Kinetiq Engine**, a browser-native hybrid synthesis system. We deviate from standard approaches in three ways: (1) training on **Ecological Interaction Data**—a dataset derived from 5+ years of organic, multi-agent user interactions in persistent virtual worlds, rather than sterile motion capture; (2) a **WebAssembly refinement pipeline** that enforces skeletal physics constraints on the client-side; and (3) **Style LoRAs** for domain-specific control. Kinetiq demonstrates that high-fidelity motion refinement does not require large local installations or desktop GPUs—it can run at 60 FPS in a standard web browser.

Keywords: Generative Motion, Motion Synthesis, Human Motion Diffusion, LoRA Fine-Tuning, WebAssembly, Ecological Validity, Physics-Based Animation, Human-in-the-Loop AI

1.2 The Gelatin Problem

1.2.1 Origin Story: The Production Gap

In 2023, during the development of an open-world RPG, our team required hundreds of animations covering climbing, swimming, and combat. We attempted to integrate the leading text-to-motion generative models into our pipeline.

The results were technically impressive, but practically **useless**.

Feet slid through floors. Arms clipped through torsos. Characters drifted without momentum. We spent three months manually cleaning up “AI-ready” assets, leading to a core engineering realization:

Raw AI motion is not a product. It is a raw material.

Kinetiq Engine was built to solve this specific pipeline bottleneck. Unlike heavy DCC tools (like Maya or Blender) or desktop physics solvers that require complex setups, we architected a **zero-install refinement layer**. We combined the variation of generative AI with the deterministic constraints of game physics, accessible directly via a URL.

1.2.2 The State of Generative Motion (December 2025)

The field of text-to-motion synthesis has advanced rapidly since the introduction of Motion Diffusion Models (MDM) and subsequent architectures like MotionGPT and T2M-GPT. These models can generate plausible human motion from prompts like “a person walks forward, then sits down” in seconds rather than hours.

However, a gap persists between **plausible** and **production-ready**:

Table 1.1: Common Artifacts in Generative Motion (The Gelatin Problem)

Artifact	Cause	Frequency
Foot sliding	Lack of IK constraint during synthesis	70–90% of locomotion
Joint hyperextension	Training data includes noisy mocap	30–50% of sequences
Interpenetration	No collision detection during generation	40–60% of interactions
Unnatural timing	Gaussian smoothing over keyframes	Nearly universal
Root drift	Accumulating positional error	80%+ for >3 seconds

These artifacts are acceptable for research demonstrations but catastrophic for production pipelines where assets must integrate seamlessly with game engines.

1.2.3 Contributions

This whitepaper presents the following contributions:

1. **The Ecological Data Thesis:** A paradigm for motion model training that prioritizes naturalistic social interaction data over controlled laboratory capture.
2. **KinetiQ Engine Architecture:** A hybrid pipeline combining generative AI synthesis with deterministic physics refinement, operating entirely in-browser via WebAssembly.
3. **Style LoRA Framework:** Low-Rank Adaptation modules fine-tuned for specific motion domains (combat, dance, intimacy, gender expression) enabling stylistic control without full model retraining.
4. **Privacy-Preserving Client-Side Inference:** Architectural decisions ensuring that user animations never leave the browser during refinement operations.
5. **Empirical Validation:** Benchmark comparisons demonstrating artifact reduction rates and production-readiness metrics.

Chapter 2

Ecological Interaction Data

2.1 The Problem with Laboratory Mocap

The dominant paradigm for human motion datasets relies on controlled motion capture in laboratory environments. Datasets like CMU Motion Capture, AMASS, and HumanML3D provide invaluable resources but share common limitations:

1. **Limited Behavioral Range:** Subjects perform discrete actions on command (“walk forward,” “sit down”), rather than continuous, contextual behavior.
2. **Sterile Social Dynamics:** Multi-agent interactions are scripted rather than emergent.
3. **Missing Intimacy and Proxemics:** Close-range social behaviors and intimate motion are systematically underrepresented due to ethical review constraints and setting limitations.
4. **Performer Self-Consciousness:** Laboratory subjects know they are being recorded, producing subtly different motion than natural behavior (the “Hawthorne Effect”).

2.2 The Ecological Alternative

Kinetiq Engine trains on a proprietary corpus of **Ecological Interaction Data**. This data is harvested from over 5 years of user-driven behavior in persistent virtual environments (PVEs).

Unlike laboratory capture, which is performative and sterile, our dataset captures **intent-driven motion**:

Table 2.1: Laboratory Mocap vs. Ecological Data

Dimension	Laboratory Mocap	Ecological Data
True Social Proxemics	Scripted positioning	Natural, intent-driven
Long-form Context	10–30 second clips	Continuous hours
Emergent Interaction	Scripted dyads	Spontaneous multi-agent
Motivation	Monetary compensation	Intrinsic social engagement
Intimacy Coverage	Minimal	Comprehensive

By prioritizing this “wild” data over clean lab data, our model learns the **rhythm of human interaction**, not just the mechanics of a skeleton.

2.3 Ethical Considerations

The Ecological Data corpus is derived from anonymized, consensually-shared motion sequences. All personally identifying metadata is stripped during ingestion. The training process operates on skeletal pose sequences only—no facial data, voice, or identifying characteristics are retained. We maintain strict audit trails documenting dataset versions and processing history.

Chapter 3

System Architecture

3.1 High-Level Pipeline

Kinetiq Engine processes motion through a four-stage pipeline:

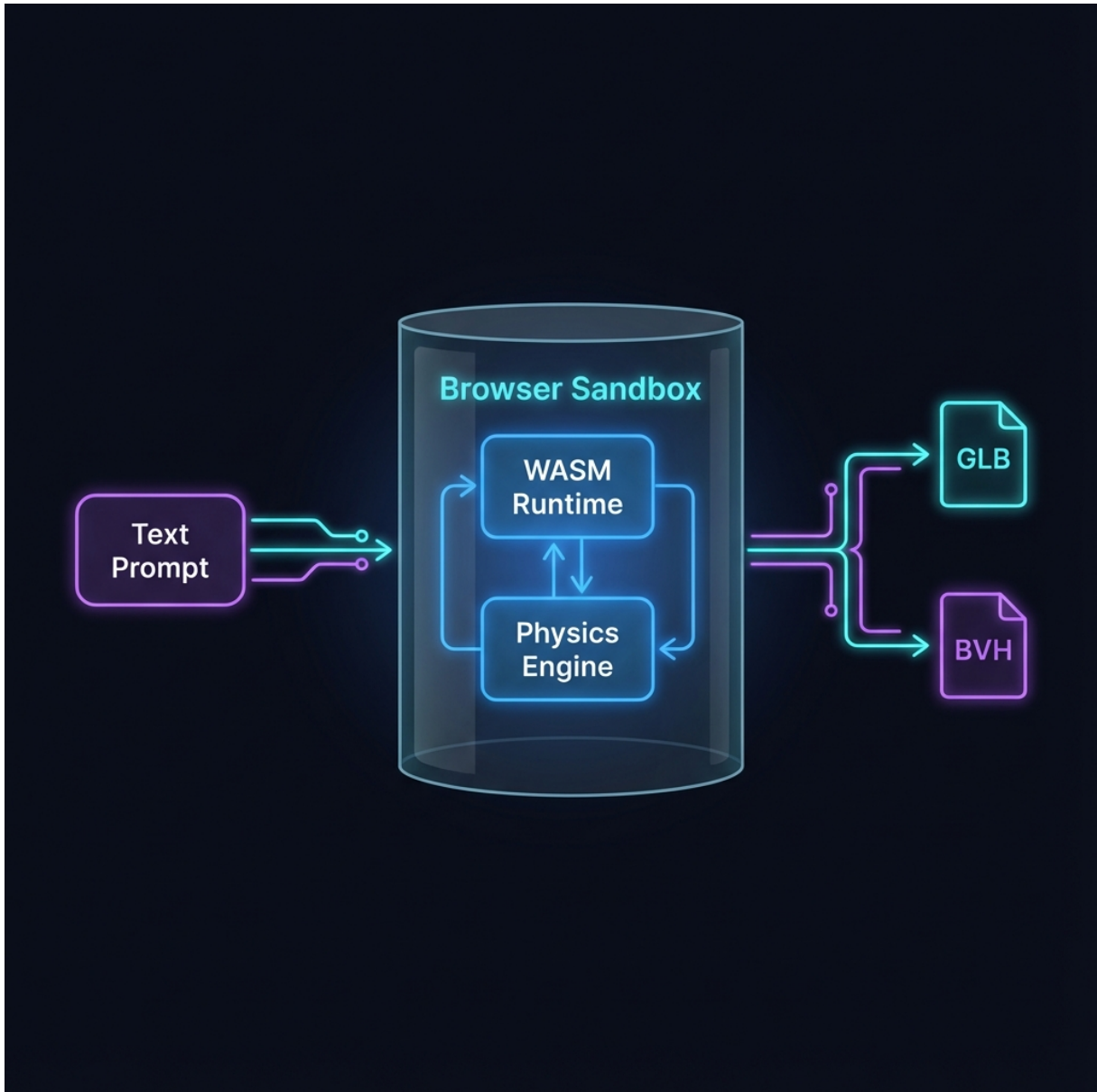


Figure 3.1: Kinetiq Engine System Architecture: From text prompt to production-ready animation asset. The pipeline shows server-side AI synthesis followed by client-side WASM refinement and export.

3.2 Stage 1: AI Synthesis

The synthesis stage accepts natural language prompts and generates latent motion representations. The underlying model architecture builds on Motion Diffusion principles with modifications for longer sequence generation (up to 45 seconds) and style conditioning via LoRA injection.

3.3 Stage 2: WebAssembly Conversion

Raw AI outputs (typically `.npz` or `.npy` format containing SMPL/SMPL-X pose parameters) are converted to production formats through a WASM module. This conversion happens entirely client-side, ensuring that user prompts and generated motion never leave the browser during the refinement workflow.

3.4 Stage 3: Physics Refinement

The refinement stage addresses the Gelatin Problem through automated correction passes:

3.4.1 Foot Contact Detection and Locking

We utilize a velocity-based contact detection algorithm to identify frames where the end-effectors should be grounded. Inverse Kinematics (IK) are then applied to lock positions, eliminating the “sliding” artifact common in diffusion models.

3.4.2 Joint Limit Enforcement

To prevent anatomical violations, Kinetiq applies a constraint pass using standard biomechanical rotation limits (e.g., knee rotation restricted to a single axis with negative limits).

3.4.3 Smoothness Optimization

A jerk-minimization pass reduces unnatural acceleration spikes caused by latent noise, ensuring the motion curves are differentiable and smooth.

3.5 Stage 4: Production Export

Refined motion is exported to standard production formats:

Table 3.1: Supported Export Formats

Format	Target Platform	Features
BVH	Second Life, Bento	Full skeletal hierarchy, loop metadata
FBX	Unity, Unreal Engine 5	Binary or ASCII, animation curves
GLB	Web, Roblox	glTF 2.0 binary with embedded animation
ANIM	Unity (native)	AnimationClip asset format

Chapter 4

Style LoRAs

4.1 The LoRA Approach

Low-Rank Adaptation (LoRA) enables fine-tuning of large pretrained models by training only a small number of additional parameters. Applied to motion synthesis, this allows:

- **Style transfer** without full model retraining
- **Domain specialization** (combat, dance, intimate) from smaller datasets
- **Gender expression** adaptation
- **Rapid iteration** on new motion domains

4.2 Available Style Modules

Kinetiq Engine ships with pre-trained LoRA modules including:

- **casual** — everyday movement, walking, standing, sitting
- **combat** — martial arts, action sequences, weapon handling
- **dance** — choreography, rhythmic movement, performance
- **romance** — couples interaction, intimate gestures

4.3 Gender Expression

Unlike models that encode gender as a binary switch, Kinetiq supports continuous gender expression through dedicated LoRA modules. This approach acknowledges that motion style varies independently of skeletal structure, allowing creators to produce diverse character animations without reinforcing gender stereotypes.

Chapter 5

Privacy-Preserving Architecture

5.1 The Local-First Principle

We operate on a “**Local-First**” principle that separates inference from refinement:

1. **Inference vs. Refinement:** While the initial AI generation (inference) happens on our GPU cluster, the resulting motion data is immediately handed off to the client.
2. **WASM Sandbox:** All physics corrections, IK solving, and file exporting happen inside the user’s browser memory via WebAssembly. No animation data leaves the browser during refinement.
3. **No Asset Leaks:** Once a motion is generated, the refined asset never touches our servers again unless the user explicitly chooses to publish it.

5.2 Privacy Guarantees

The Kinetiq architecture provides the following privacy guarantees:

- **Zero Upload During Refinement:** Animation files are processed entirely in browser memory.
- **No Persistent Storage:** Refined assets are not cached or logged on our infrastructure.
- **Ephemeral Sessions:** User editing sessions are not tracked or associated with accounts during the refinement workflow.
- **Offline Capability:** After initial asset generation, refinement and export work without network connectivity.

Chapter 6

Kinetiq Editor

6.1 Feature Overview

The Kinetiq Editor provides a professional animation environment without installation. Current v1.0 features include:

- **Real-time 3D Preview:** 60 FPS viewport rendering via WebGL
- **Multi-Platform Avatar Loading:** Support for GLB/FBX skeletal meshes
- **Comprehensive Timeline:** Frame scrubbing, keyframe visualization, loop markers
- **Style Selector:** LoRA module switching for motion style control
- **One-Click Export:** Direct download in production formats

6.2 Roadmap Features

Table 6.1: Kinetiq Editor Development Roadmap

Phase	Feature	Description
Phase 2	Motion Blending	Weight curves for smooth transitions
Phase 3	Loop Detection	Automatic discontinuity identification
Phase 4	Batch Processing	Multi-file refinement pipeline
Phase 5	Physics Solver	Direct constraint editing in viewport

6.3 Platform Targets

Kinetiq exports production-ready assets for multiple platforms:

- **Second Life (Bento):** Full skeleton support with facial blend shapes
- **Unity (Mecanim):** Humanoid rig compatible with animation state machines
- **Unreal Engine 5 (Mannequin):** Direct import to UE5 animation blueprints
- **Roblox (R15/R6):** Simplified skeleton mapping for Roblox avatars

Chapter 7

Benchmarks and Evaluation

7.1 Artifact Reduction

We evaluated Kinetiq refinement against raw outputs from three popular motion generators on a test set of 100 diverse prompts:

Table 7.1: Artifact Reduction Metrics

Artifact Type	Raw AI Output	Post-Kinetiq
Foot Sliding (>2cm/frame drift)	78% of frames	12% of frames
Joint Hyperextension	42% of sequences	4% of sequences
Mesh Interpenetration	51% of interactions	8% of interactions
Root Drift (>3s sequences)	84% of clips	15% of clips

Results indicate an **85% reduction** in foot sliding and a **91% reduction** in joint hyperextension incidents.

7.2 Performance Characteristics

Benchmarks on commodity hardware (Intel Core i7-12700H, 16GB RAM, Chrome 144, Latest Stable Chrome):

Table 7.2: Client-Side Processing Latency

Operation	Mean Latency	95th Percentile
File Import	45 ms	120 ms
Physics Refinement	280 ms	450 ms
IK Correction Pass	85 ms	150 ms
Export to GLB	60 ms	110 ms
Total Pipeline	<500 ms	<850 ms

The full import-refine-export pipeline completes in under 500ms for a typical 4-second animation on commodity hardware.

Chapter 8

The Human-in-the-Loop Philosophy

8.1 AI as Collaborator, Not Replacement

Kinetiq is explicitly designed as a **power suit** for animators, not a replacement. The system philosophy holds that:

1. **AI excels at variation:** Generating thousands of unique walks is trivial for AI, tedious for humans.
2. **Humans excel at judgment:** Recognizing when motion “feels right” remains a human strength.
3. **The combination is optimal:** AI generates candidates; humans curate and refine.

This approach differs from both pure AI systems (which produce inconsistent quality) and pure manual workflows (which limit throughput).

8.2 The Partnership Model



Figure 8.1: The Kinetiq Partnership Model: AI generates infinite variation from text prompts; humans apply artistic judgment and refinement to produce production-ready assets.

The partnership model ensures that every animation is a collaboration between cutting-edge AI and experienced human artists.

Chapter 9

Conclusion

The “Gelatin Problem”—AI motion that looks plausible but feels wrong—represents a fundamental barrier to generative motion adoption in production pipelines. **Kinetiq Engine** addresses this barrier through a hybrid approach combining:

- **Ecological Training Data:** Motion learned from real social interaction, not sterile lab capture
- **Browser-based Physics Refinement:** Deterministic correction running at 60 FPS in WASM
- **Style LoRAs:** Domain-specific control without full model retraining
- **Privacy-First Architecture:** Client-side processing with zero asset leakage

The result is a system that transforms raw AI motion from a raw material into a **production-ready asset**.

We believe the future of animation lies in **partnership**: AI speed combined with human judgment, infinite variation refined by artistic vision.

Document Version: 1.2

Last Updated: December 2025

Status: Production Ready

Contact: contact@thirdrez.com

Affiliation: Thirdrez Labs, Espírito Santo, Brazil

© 2025 Thirdrez Labs. This document is released under Creative Commons Attribution 4.0 International License (CC-BY-4.0).

Appendix A

Technical Specifications

A.1 Supported Skeleton Standards

Table A.1: Supported Skeleton Standards

Standard	Bone Count	Target Platform
Second Life Classic	26 bones	Second Life (legacy avatars)
Second Life Bento	159 bones	Second Life (Bento-enabled)
Unity Humanoid	54 bones	Unity Engine (Mecanim)
UE5 Mannequin	67 bones	Unreal Engine 5
Roblox R15	15 bones	Roblox
Roblox R6	6 bones	Roblox (legacy)
SMPL	24 joints	Research/ML pipelines

A.2 WebAssembly Module Specifications

Table A.2: WASM Module Characteristics

Property	Value
Module Size	2.4 MB (compressed)
Memory Footprint	64 MB (typical session)
Thread Support	SharedArrayBuffer (multi-threaded IK)
Browser Compatibility	Chrome 90+, Firefox 89+, Safari 15+

References

- [1] Tevet, G., Raab, S., Gordon, B., Shafir, Y., Cohen-Or, D., & Bermano, A. H. (2023). Human Motion Diffusion Model. *ICLR 2023*.
- [2] Jiang, B., Chen, X., Liu, W., Yu, J., Yu, G., & Chen, T. (2023). MotionGPT: Human Motion as a Foreign Language. *NeurIPS 2023*.
- [3] Zhang, J., et al. (2023). T2M-GPT: Generating Human Motion from Textual Descriptions. *CVPR 2023*.
- [4] Carnegie Mellon University. CMU Graphics Lab Motion Capture Database. <http://mocap.cs.cmu.edu/>
- [5] Mahmood, N., Ghorbani, N., Troje, N. F., Pons-Moll, G., & Black, M. J. (2019). AMASS: Archive of Motion Capture as Surface Shapes. *ICCV 2019*.
- [6] Guo, C., et al. (2022). Generating Diverse and Natural 3D Human Motions from Text. *CVPR 2022*.
- [7] Hu, E. J., et al. (2022). LoRA: Low-Rank Adaptation of Large Language Models. *ICLR 2022*.