Acht Mac Minis.
Ein lernendes System.

Wie aus alter Intel-iGPU-Hardware, sichtbaren Flaschenhälsen und einer kompromisslosen Beweiskette ein eigener Modellbau-Windkanal entstand.

8	8,5 Mio.	17
REGISTRIERTE IGPU-WORKER	PARAMETER	VERSIEGELTE SCHRITTE

Zwischenbericht, keine klassische Forschungsarbeit. Eine technische Biografie für Menschen, die verstehen wollen, wie ein ungewöhnliches System entsteht - einschließlich der Umwege.

Joachim Müller-Peter / Unabhängige Forschung / itbois / Schweiz

Ein System beginnt zu lernen

DIE MELDUNG IN EINEM SATZ

Acht alte Mac Minis haben einen eigenen Transformer gemeinsam trainiert und auf allen acht Intel-iGPUs denselben versiegelten Modellstand erzeugt.

Das Ergebnis ist noch kein fertiges Mathematik- oder Coding-Modell. Es ist der belastbare Beweis, dass der komplette Weg funktioniert: Daten, Tokenizer, Forward-Pass, Backpropagation, Optimizer, Netzwerk, Temperaturregeln, Reproduzierbarkeit und öffentliche Dokumentation.

5,98 %	190,38	0
VALIDATION-LOSS REDUZIERT	GÜLTIGE TOKEN/S LOKAL	CPU-FALLBACKS

WAS BEREITS STEHT

- **ECHTES TRAINING**
  17 semantische AdamW-Schritte
- **VERTEILT**
  acht lokale Gradienten, ein gemeinsamer Update
- **REPRODUZIERBAR**
  identische Modell- und Optimizer-Hashes
- **ÖFFENTLICH**
  Read-only-Takt auf itbois.ch

WAS WIR NICHT BEHAUPTEN

- **KEINE FÄHIGKEIT**
  noch kein privater Fähigkeitstest
- **KEINE VOLLE EPOCHE**
  bisher rund 4,3 % eines Token-Durchlaufs
- **KEIN WELTWISSEN**
  nur freigegebene Mathematik- und Coding-Daten
- **KEIN PAPER**
  bewusst eine technische Zwischenbilanz

Die eigentliche Nachricht: Der Windkanal ist nicht länger nur eine Idee oder Oberfläche. Er ist eine funktionierende, messende und veröffentlichende Modellbau-Infrastruktur.

Vom Vertrag zum Schritt 17

Nicht der direkte Weg ist die Geschichte. Die belastbare Kette aus Entscheidungen, Messungen und Korrekturen ist es.

02. SEP
MOMENT NULL
From-scratch, Mathematik, Coding und Evidenzgrenzen werden festgeschrieben.

02. SEP
ACHT GPUs ANSPRECHBAR
Alle Worker bestehen den OpenCL-Pfad ohne CPU-Fallback.

02. SEP
DATEN VERSIEGELT
3.459 kontrollierte Datensätze, eigener 8.192er-Tokenizer, keine Eval-Leckage.

03. SEP
ERSTER LERNSIGNAL-PILOT
16 echte GPU-Schritte auf node8. Validation-Loss sinkt um 5,98 Prozent.

03. SEP
ACHT-GPU-VOLLFORMAUFBAU
Synthetische Vollform zeigt: Berechnung, Transfer und Hashgleichheit funktionieren.

03. SEP
FEHLER SICHTBAR GEMACHT
Last-Holds, Temperatur-Holds und ein Protokollfehler werden nicht versteckt, sondern versiegelt.

03. SEP
SCHRITT 17
Acht reale Gradienten, node6-iGPU-Mittelung, acht identische Checkpoints.

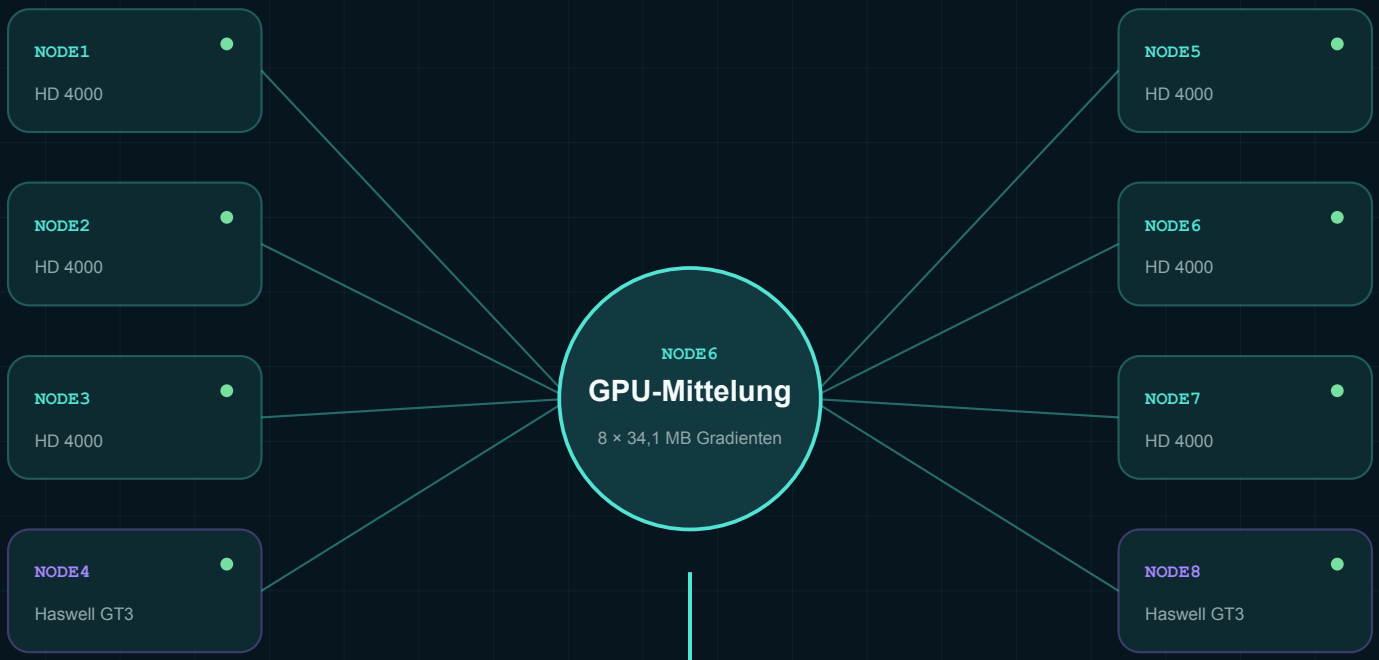
04. SEP
ZWISCHENBERICHT
Die technische Biografie wird als eigenständige Publikation geöffnet.

DAS PRINZIP

Jeder Umweg hinterlässt einen überprüfbaren Beleg.

Wie acht GPUs einen Schritt schreiben

Synchrones Data-Parallel-Training auf zwei gemessenen Intel-iGPU-Kohorten.



1/8 + 1/8 + ... + 1/8

ADAMW-SCHRITT 17 AUF ALLEN ACHT INTEL-IGPUS

8 identische Modellhashes

13d9d666a303... / Optimizer e2e35509daf2...

RECHENREGEL

Modellarithmetik nur auf Intel-iGPU. CPU orchestriert, überträgt und prüft Hashes.

GEMESSENE FLOTTE

Sechs Macmini6,2 mit HD 4000; zwei Macmini7,1 mit Haswell GT3.

Klein genug zum Begreifen

WNM-1S ist bewusst kein verkleinertes Weltmodell, sondern ein fokussierter Modellbaukörper.

MODELLKÖRPER

gekoppelte Ein- und Ausgabegewichte

8.524.032 PARAMETER	8 TRANSFORMER-LAYER
256 MODELLBREITE	4 × 64 ATTENTION-HEADS
704 MLP-BREITE	256 KONTEXT-TOKENS
8.192 TOKENIZER-VOKABULAR	FP32 RECHENFORMAT

TRAININGSKORPUS / 142.534 TOKENS

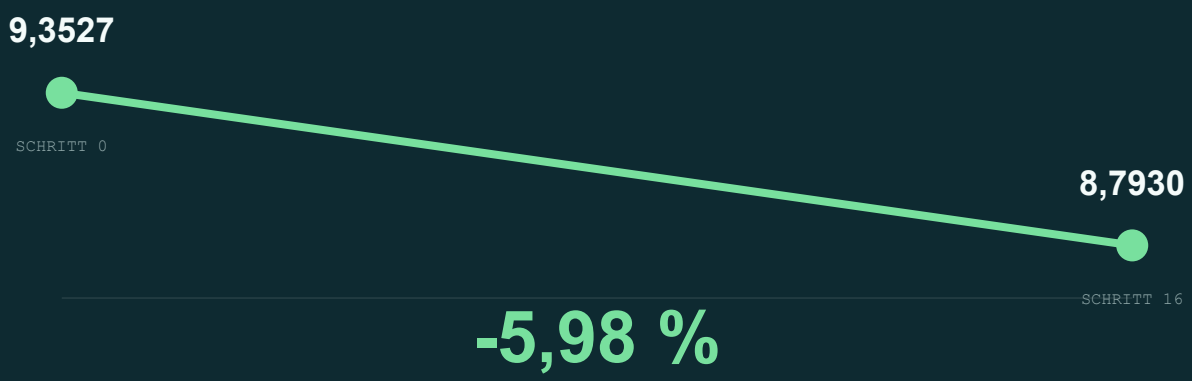
DIE POSITIVE NÄCHSTE BAUSTELLE

Die Mathematikbasis steht. Für ein wirklich Coding-orientiertes Modell wird der kontrollierte Codekorpus nun substanziell erweitert - ohne Tests, Verifier, privates Eval oder Weltwissen einzumischen.

Lernbarkeit ist sichtbar

Zwei getrennte Messarten: Validation zeigt Veränderung auf einem festen Fenster; Trainings-Loss zeigt den rauen Schrittverlauf.

FESTES VALIDIERUNGSFENSTER



-5,98 %

UNGEGLÄTTETER TRAININGS-LOSS / SCHRITTE 1 BIS 17

Rohwerte dürfen schwanken

SCHRITT 17 / 8-NODE-MITTEL 8,0044

Lernbarkeit: ja. Fähigkeit: noch offen.

Der fallende Validation-Loss ist ein echtes Signal auf genau einem festgelegten Fenster. Erst die getrennte private Evaluation kann zeigen, ob daraus Mathematik- oder Coding-Fähigkeit entstanden ist.

Der Flaschenhals wurde zum Fenster

Der Lauf dauerte 449,46 Sekunden. Die meiste Zeit lag nicht in der Arithmetik, sondern im Eintritt und im Transport.

7 min 29 s

GESAMTER VERTRAGLICHER RHYTHMUS

Eintritt + Prep	152,95 s	**Inputs**	142,13 s
lokale GPU	36,64 s	**Gradienten holen**	44,85 s
GPU-Mittelung	9,58 s	**Gradient verteilen**	46,06 s
AdamW + Seal	7,55 s		

Hinweis: Die 149,14 s node4-Wartezeit sind im Eintritts-/Prep-Block enthalten.

846 MB

PRIVATE INPUTS

Checkpoint, Optimizer, Tokenizer und Trainingssplit an acht Worker.

273 MB

GRADIENTEN HIN

Acht lokale Gradientenzustände wandern zu node6.

273 MB

GRADIENT ZURÜCK

Der gemittelte Zustand wird an alle Worker zurückverteilt.

DIE MARKETING-WAHRHEIT IST AUCH DIE TECHNISCHE WAHRHEIT

Der Netzwerkengpass wird nicht kaschiert. Er zerlegt das Training in sichtbare Akte: warten, verteilen, rechnen, sammeln, mitteln, zurücksenden, versiegeln. Genau deshalb kann das Publikum dem Modell beim Entstehen zusehen.

Worauf man heute stolz sein darf

01

Von Grund auf

Kein fertiges Basismodell wurde übernommen. Architektur, Tokenizer, Datenweg und Gewichte wurden im Projekt festgelegt.

02

GPU-only bewiesen

Forward, Backward, Gradientenmittelung und AdamW liefen ohne CPU-Modellrechnung auf alter Intel-iGPU-Hardware.

03

Fehler wurden Teil der Beweiskette

Thermal-Holds, Last-Holds und der Protokollfehler sind sichtbar geblieben und haben die Strecke verbessert.

04

Acht gleiche Endzustände

Der gemeinsame Schritt endete nicht nur erfolgreich, sondern mit identischen Modell- und Optimizer-Hashes.

05

Öffentlich, aber nicht offen steuerbar

Besucher sehen Zustände und Messwerte. Gewichte, Trainingsdaten und Steuerung bleiben privat.

Das wertvollste Ergebnis ist noch nicht die Intelligenz des Modells.
Es ist der entstandene Modellbau-Windkanal: reproduzierbar, sichtbar, erweiterbar und ehrlich über seine Grenzen.

Die eigentliche Forschung beginnt

Die Infrastruktur hat bestanden. Nun muss aus einem Lernsignal belastbare Domänenfähigkeit werden.

GATE 01 / HARDWAREGRENZE KLÄREN

Strikt 2012 oder bewusst zwei iGPU-Kohorten?

Die gemessene Flotte umfasst sechs Macmini6,2/HD 4000 und zwei Macmini7,1/Haswell GT3. Für den nächsten Modellzweig wird diese Grenze ausdrücklich entschieden.

02

Schritt 17 privat evaluieren

Gewichte unverändert lassen. Festes Validation-Fenster und 256 getrennte Aufgaben messen.

03

Coding-Daten ausbauen

Die drei heutigen Coding-Einträge zu einem kontrollierten, lizenzierten Domänenkorpus erweitern.

04

Eine echte Epoche definieren

Mit acht Workern ungefähr 70 synchrone globale Schritte über den heutigen Tokenumfang.

05

Netzwerk amortisieren

Korpus und Checkpoint resident halten; längere versiegelte Segmente gegen den Baseline-Pfad testen.

DER NÄCHSTE BELASTBARE SATZ

Nicht: Das Modell ist schon gut. Sondern: Wir können jetzt präzise messen, ob und wie es gut wird.

Menschen, Projekt, Kontakt

ÜBER JOACHIM MÜLLER-PETER

Technik wird interessant, wenn sie erklärbar wird.

Joachim Müller-Peter arbeitet an der Schnittstelle von gewachsenen KMU-Systemen, Datenbanken, Web-Anwendungen, FileMaker, lokaler Infrastruktur und künstlicher Intelligenz. Im Model Windkanal behandelt er Modelle als vollständige technische Systeme - mit Hardware, Energie, Netzen, Regeln, Messgrenzen und Verantwortung.

KURZPROFIL WNM-1S

MODELL	**8,5 Mio. Parameter**
DOMÄNE	**Mathematik + Coding**
WELTWISSEN	**bewusst ausgeschlossen**
STAND	**Schritt 17 versiegelt**
CLAIM	**Lernbarkeit, nicht Fähigkeit**

KONTAKT

Joachim Müller-Peter

Unabhängige Forschung · itbois · Schweiz

contact@itbois.ch

linkedin.com/in/joachim-müller-5201242b2

ÖFFENTLICHES SCHAUFENSTER

itbois.ch/apps/windkanal-laboratory/

Live-Takt, versiegelte Meilensteine, Loss-Messungen, Veröffentlichungen und die acht Worker - read-only und ohne Zugriff auf Trainingsdaten oder Gewichte.

EVIDENZSTAND

Dieser Zwischenbericht stützt sich auf die versiegelten Projektbelege WNM-0044, WNM-0045, WNM-0054, WNM-0064 und WNM-0070 bis WNM-0076. Er ist eine journalistisch-technische Einordnung und kein Peer-Review-Paper. Messwerte gelten für die beschriebenen Läufe und Hardwarezustände.

Eight Mac minis.
One learning system.

How ageing Intel integrated GPUs, visible bottlenecks and an uncompromising evidence chain became a model-building wind tunnel.

8	8.5m	17
REGISTERED IGPU WORKERS	PARAMETERS	SEALED STEPS

An interim report, not a conventional research paper. A technical biography for people who want to understand how an unusual system is built - detours included.

Joachim Müller-Peter / Independent research / itbois / Switzerland

A system begins to learn

THE STORY IN ONE SENTENCE

Eight old Mac minis have jointly trained a purpose-built Transformer and produced the same sealed model state on all eight Intel iGPUs.

The result is not yet a finished mathematics or coding model. It is robust proof that the full path works: data, tokenizer, forward pass, backpropagation, optimizer, network, thermal rules, reproducibility and public documentation.

5.98%	190.38	0
VALIDATION LOSS REDUCTION	AGGREGATE VALID TOKEN/S	CPU FALLBACKS

WHAT ALREADY EXISTS

- **REAL TRAINING**
  17 semantic AdamW steps
- **DISTRIBUTED**
  eight local gradients, one update
- **REPRODUCIBLE**
  identical model and optimizer hashes
- **PUBLIC**
  read-only pulse on itbois.ch

WHAT WE DO NOT CLAIM

- **NO CAPABILITY CLAIM**
  no private capability test yet
- **NO FULL EPOCH**
  about 4.3% of one token pass
- **NO WORLD KNOWLEDGE**
  approved mathematics and code only
- **NOT A PAPER**
  deliberately an interim report

The real news: the Windkanal is no longer just an idea or an interface. It is a functioning model-building infrastructure that measures and publishes its own development.

From the contract to step 17

The story is not a straight path. It is the durable chain of decisions, measurements and corrections.

02 SEP
MOMENT ZERO
From-scratch, mathematics, coding and evidence boundaries are fixed.

02 SEP
EIGHT GPUS ADDRESSABLE
Every worker passes the OpenCL path without CPU fallback.

02 SEP
DATA SEALED
3,459 controlled records, a custom 8,192-token vocabulary and no eval leakage.

03 SEP
FIRST LEARNING SIGNAL
Sixteen real GPU steps on node8. Validation loss falls by 5.98 percent.

03 SEP
EIGHT-GPU DRESS REHEARSAL
A synthetic full-form run proves compute, transport and hash equality.

03 SEP
FAILURES MADE VISIBLE
Load holds, thermal holds and a protocol error are sealed, not hidden.

03 SEP
STEP 17
Eight real gradients, node6 iGPU averaging, eight identical checkpoints.

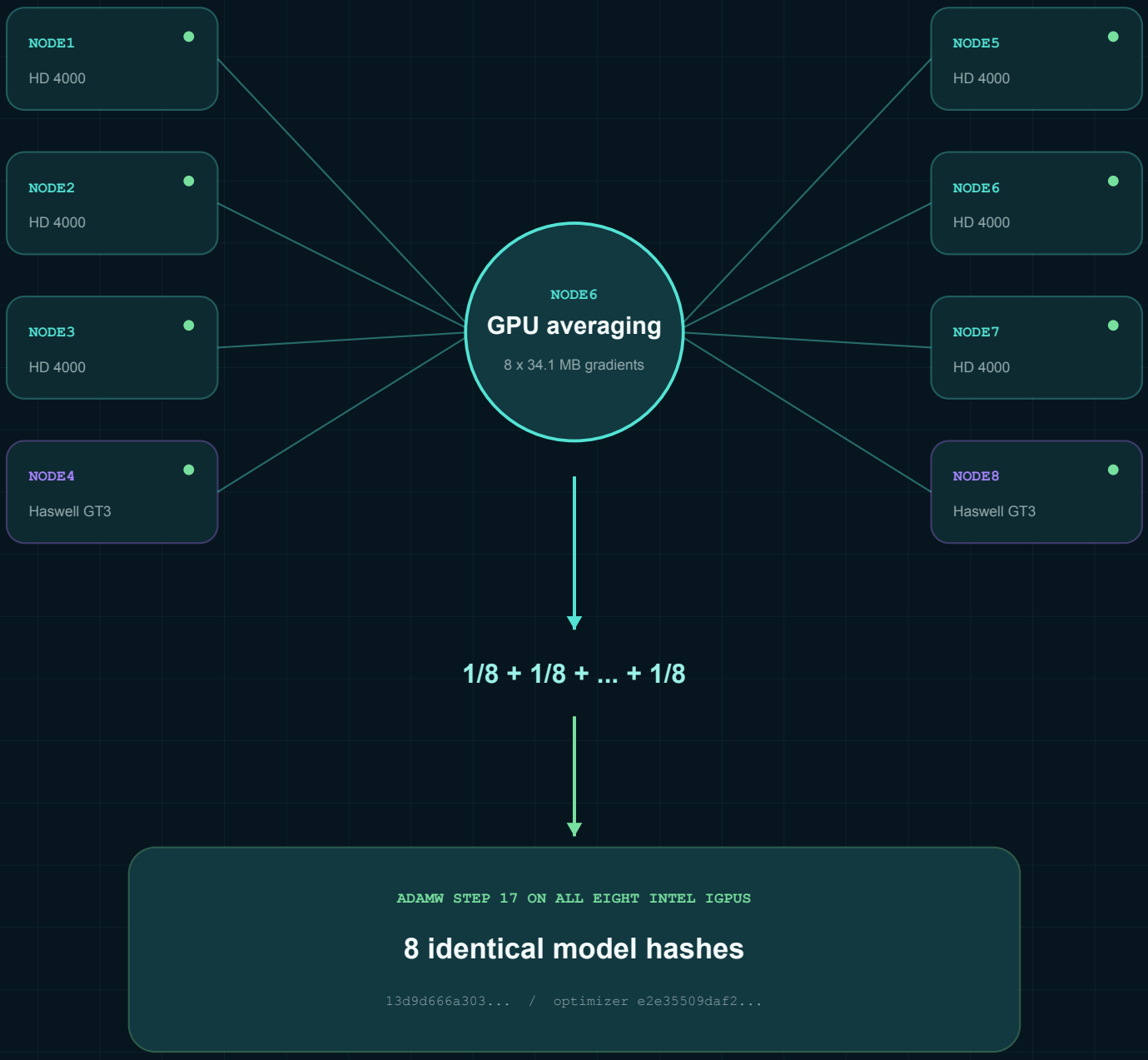
04 SEP
INTERIM REPORT
The technical biography opens as a publication of its own.

THE PRINCIPLE

Every detour leaves behind verifiable evidence.

How eight GPUs write one step

Synchronous data-parallel training across two measured Intel-iGPU cohorts.



COMPUTE RULE

Model arithmetic runs only on Intel iGPU. The CPU orchestrates, transports and verifies hashes.

MEASURED FLEET

Six Macmini6,2 systems with HD 4000; two Macmini7,1 systems with Haswell GT3.

Small enough to understand

WNM-1S is not a miniature world model. It is a deliberately focused model-building body.

MODEL BODY

LAYER 08 / ATTENTION + SWIGLU + RMS NORM

LAYER 07 / ATTENTION + SWIGLU + RMS NORM

LAYER 06 / ATTENTION + SWIGLU + RMS NORM

LAYER 05 / ATTENTION + SWIGLU + RMS NORM

LAYER 04 / ATTENTION + SWIGLU + RMS NORM

LAYER 03 / ATTENTION + SWIGLU + RMS NORM

LAYER 02 / ATTENTION + SWIGLU + RMS NORM

LAYER 01 / ATTENTION + SWIGLU + RMS NORM

tied input and output weights

8,524,032 PARAMETERS	8 TRANSFORMER LAYERS
256 MODEL WIDTH	4 x 64 ATTENTION HEADS
704 MLP WIDTH	256 CONTEXT TOKENS
8,192 TOKEN VOCABULARY	FP32 COMPUTE FORMAT

TRAINING CORPUS / 142,534 TOKENS

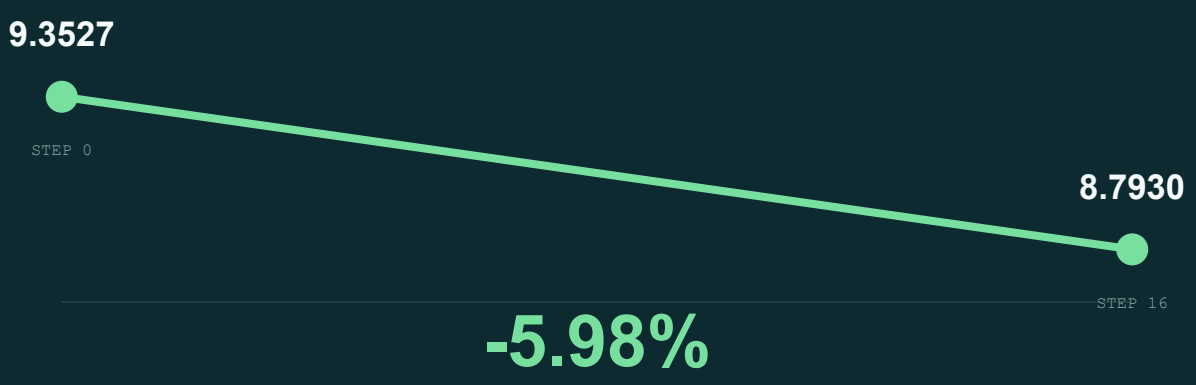
THE POSITIVE NEXT BUILDING SITE

The mathematics foundation is in place. A truly coding-oriented model now needs a substantially larger controlled code corpus - without mixing in tests, verifiers, private evaluation or world knowledge.

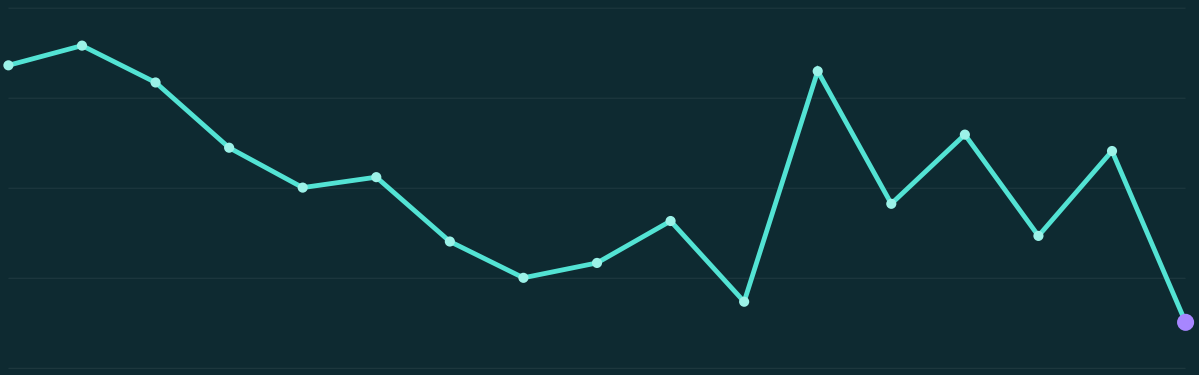
Learnability is visible

Two separate measurements: validation tracks change on a fixed window; training loss shows the raw step-by-step path.

FIXED VALIDATION WINDOW



UNSMOOTHED TRAINING LOSS / STEPS 1 TO 17



Raw values are allowed to fluctuate

STEP 17 / 8-NODE MEAN 8.0044

Learnability: yes. Capability: still open.

The lower validation loss is a genuine signal on exactly one fixed window. Only separate private evaluation can establish whether mathematics or coding capability has emerged.

The bottleneck became a window

The run took 449.46 seconds. Most time was spent not in arithmetic, but in admission and transport.

7 min 29 s

TOTAL CONTRACTUAL RHYTHM

admission + prep	152.95 s	inputs	142.13 s
local GPU	36.64 s	pull gradients	44.85 s
GPU averaging	9.58 s	push gradient	46.06 s
AdamW + seal	7.55 s		

Note: node4's 149.14 s wait is included in admission/prep.

846 MB

PRIVATE INPUTS

Checkpoint, optimizer, tokenizer and training split to eight workers.

273 MB

GRADIENTS IN

Eight local gradient states travel to node6.

273 MB

GRADIENT OUT

The averaged state is redistributed to all workers.

THE MARKETING TRUTH IS ALSO THE TECHNICAL TRUTH

The network bottleneck is not concealed. It breaks training into visible acts: wait, distribute, compute, collect, average, return and seal. That is exactly why an audience can watch a model come into being.

What is worth celebrating today

01

Built from scratch

No finished base model was adopted. The project defines the architecture, tokenizer, data path and initial weights.

02

GPU-only proven

Forward, backward, gradient averaging and AdamW ran without CPU model arithmetic on ageing Intel iGPUs.

03

Failures joined the evidence chain

Thermal holds, load holds and the protocol error remained visible and improved the system.

04

Eight identical end states

The shared step succeeded with identical model and optimizer hashes on every worker.

05

Public, never publicly steerable

Visitors see states and measurements. Weights, training data and control remain private.

The most valuable result is not yet the model's intelligence.
It is the model-building wind tunnel itself: reproducible, visible, extensible and honest about its limits.

The real research begins

The infrastructure has passed. A learning signal must now become robust domain capability.

GATE 01 / CLARIFY THE HARDWARE BOUNDARY

Strictly 2012, or deliberately two iGPU cohorts?

The measured fleet contains six Macmini6,2/HD 4000 and two Macmini7,1/Haswell GT3 systems. The next model branch must decide this boundary explicitly.

02

Privately evaluate step 17

Freeze the weights. Measure the fixed validation window and 256 separate tasks.

03

Expand coding data

Grow today's three coding records into a controlled, licensed domain corpus.

04

Define one real epoch

At today's token volume, roughly 70 synchronous global steps with eight workers.

05

Amortise the network

Keep corpus and checkpoint resident; compare longer sealed segments with the baseline path.

THE NEXT DEFENSIBLE SENTENCE

Not: the model is already good. But: we can now measure precisely whether and how it becomes good.

People, project, contact

ABOUT JOACHIM MÜLLER-PETER

Technology matters when it becomes explainable.

Joachim Müller-Peter works where long-lived SME systems, databases, web applications, FileMaker, local infrastructure and artificial intelligence meet. In the Model Windkanal, he treats models as complete technical systems - including hardware, energy, networks, rules, measurement boundaries and responsibility.

WNM-1S AT A GLANCE

MODEL	**8.5m parameters**
DOMAIN	**mathematics + coding**
WORLD KNOWLEDGE	**deliberately excluded**
STATE	**step 17 sealed**
CLAIM	**learnability, not capability**

CONTACT

Joachim Müller-Peter

Independent research · itbois · Switzerland

contact@itbois.ch

linkedin.com/in/joachim-müller-5201242b2

PUBLIC WINDOW

itbois.ch/apps/windkanal-laboratory/

Live pulse, sealed milestones, loss measurements, publications and eight workers - read-only, without access to training data or weights.

EVIDENCE STATE

This report is based on sealed project evidence WNM-0044, WNM-0045, WNM-0054, WNM-0064 and WNM-0070 through WNM-0076. It is a journalistic and technical account, not a peer-reviewed paper. Measurements apply to the runs and hardware states described here.