



Agent P-DF: Engineering a 62.5M parameter local extraction studio

Models handle geometry, deterministic code handles semantics, and UI handles the residual error.

Mohamed Sameh

Mohamed Hani

Mohamed Atta



Outlines

Introduction and Journey

The YOLO Saga & Early Setbacks

The Synthetic Dataset Odyssey

Table Detection

Table Structure Recognition

OCR & Heuristics Integration

Honorable Mentions

The Product Experience & Studio

Product As Business

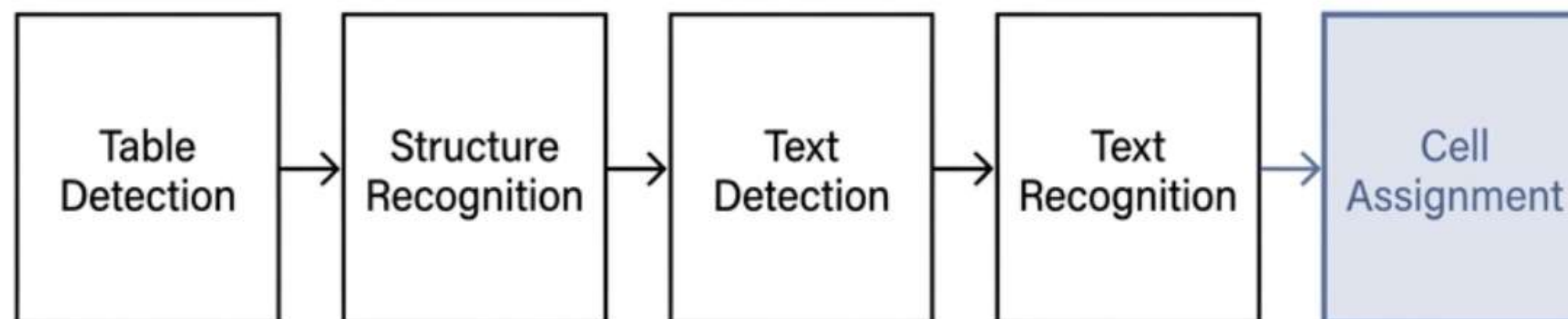
Introduction and Journey

Introduction and Journey

Architecting for the real world

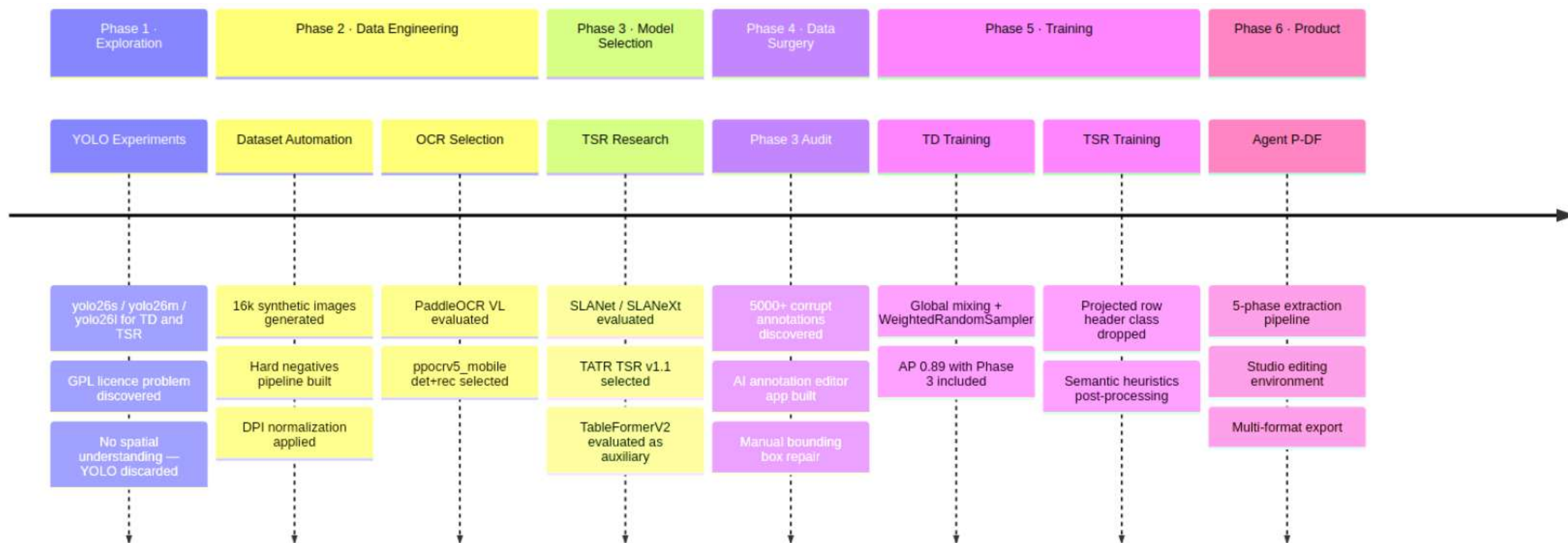
“AI is imperfect, but the editing environment should be perfect.”

Agent P-DF is built on the reality that neural networks will inevitably fail on edge cases. The system is designed not just to extract, but to empower surgical, human-in-the-loop repair.



Introduction and Journey

Timeline



The YOLO Saga & Early Setbacks

The YOLO Saga & Early Setbacks

The Setback: YOLO Architecture

Visual Texture Recognition

Predicts bounding boxes based on repetitive visual patterns (e.g., the word 'Done' in a header), resulting in zero spatial semantic understanding.

GPL Licensing

Constraint forcing unwanted open-source compliance for the entire studio.

The Requirement: DETR Architecture

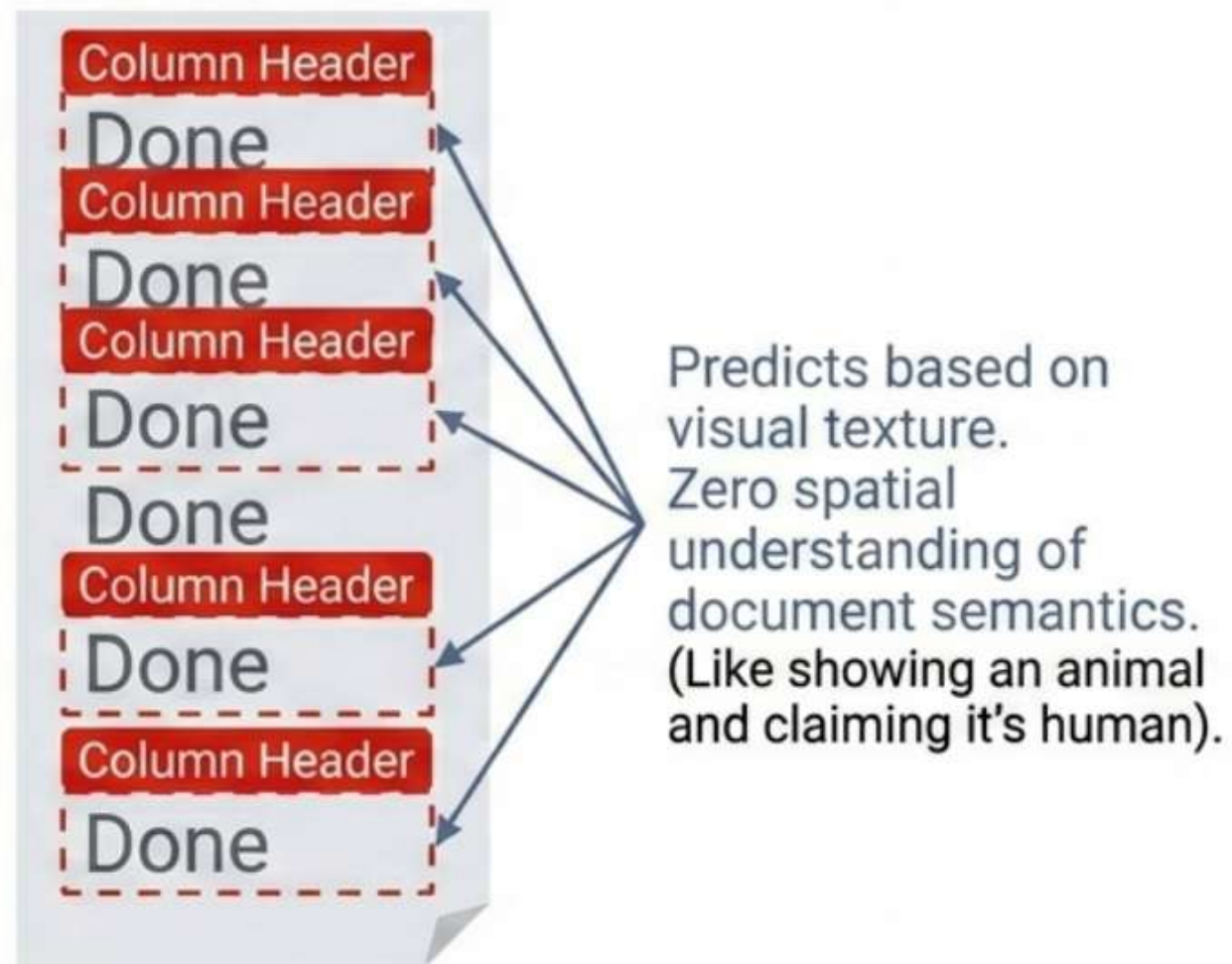
Spatial Semantic Understanding

True spatial relationship modeling via Transformer decoders.

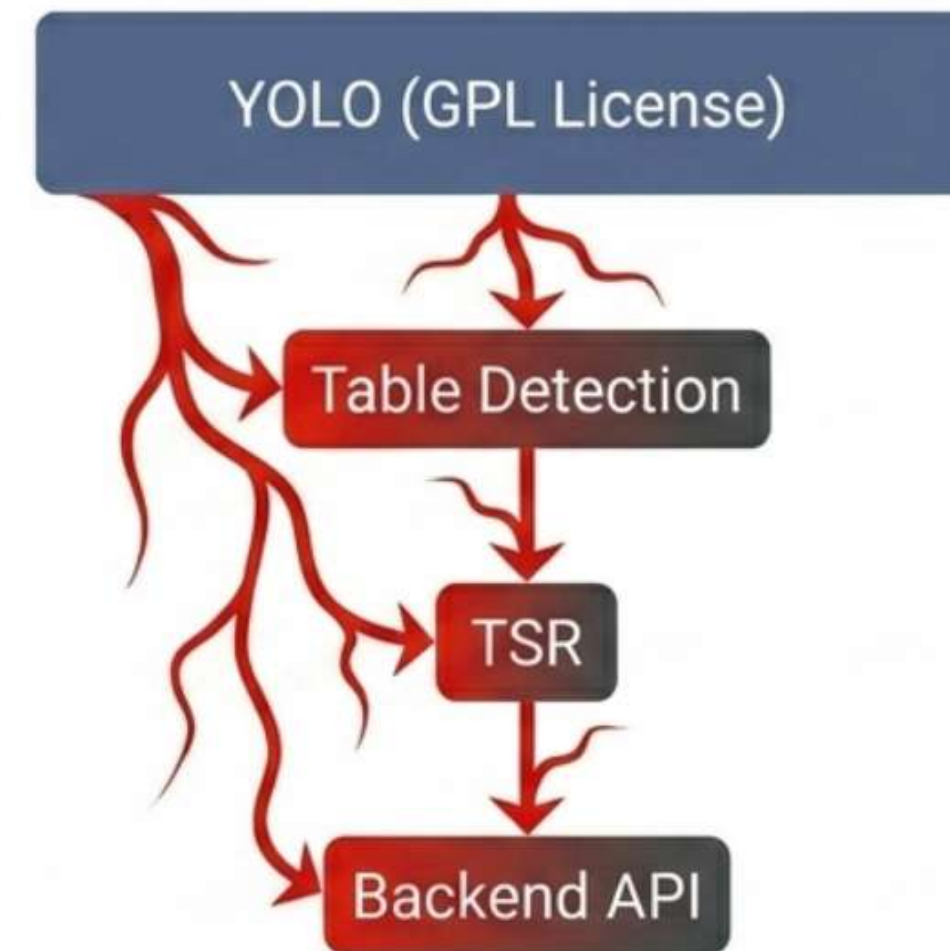
The YOLO Saga & Early Setbacks

Architecture Trap

Visual Texture vs. Spatial Semantics

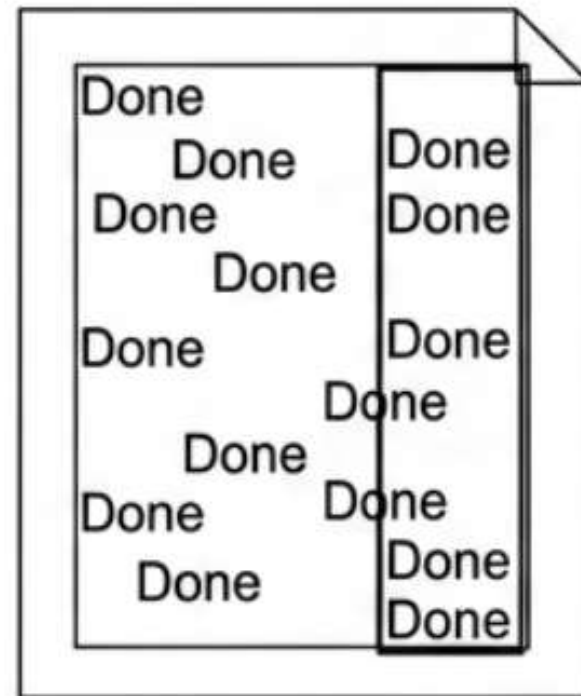


The GPL Trap



The YOLO Saga & Early Setbacks

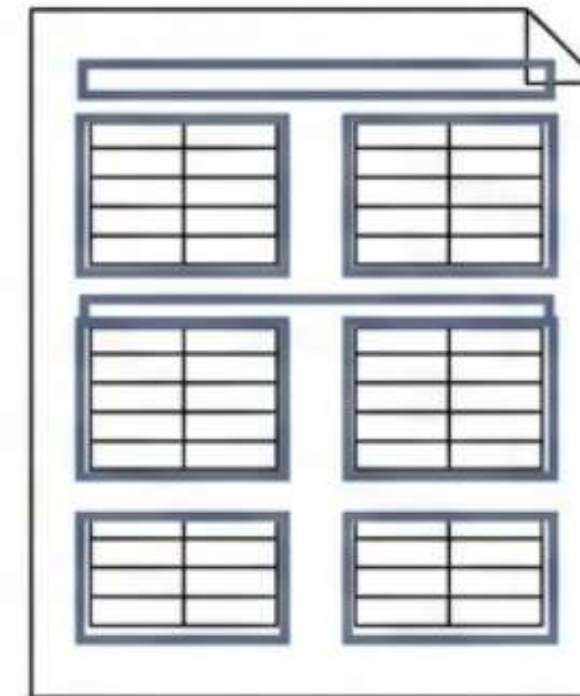
The YOLO False Start



Initial tests scaled from yolo26s to yolo26l. YOLO predicts bounding boxes based on visual texture, lacking spatial semantic understanding.

A model taught to find “texture” will flag any repeated text as a column header, regardless of surrounding grid logic.

The Semantic Shift

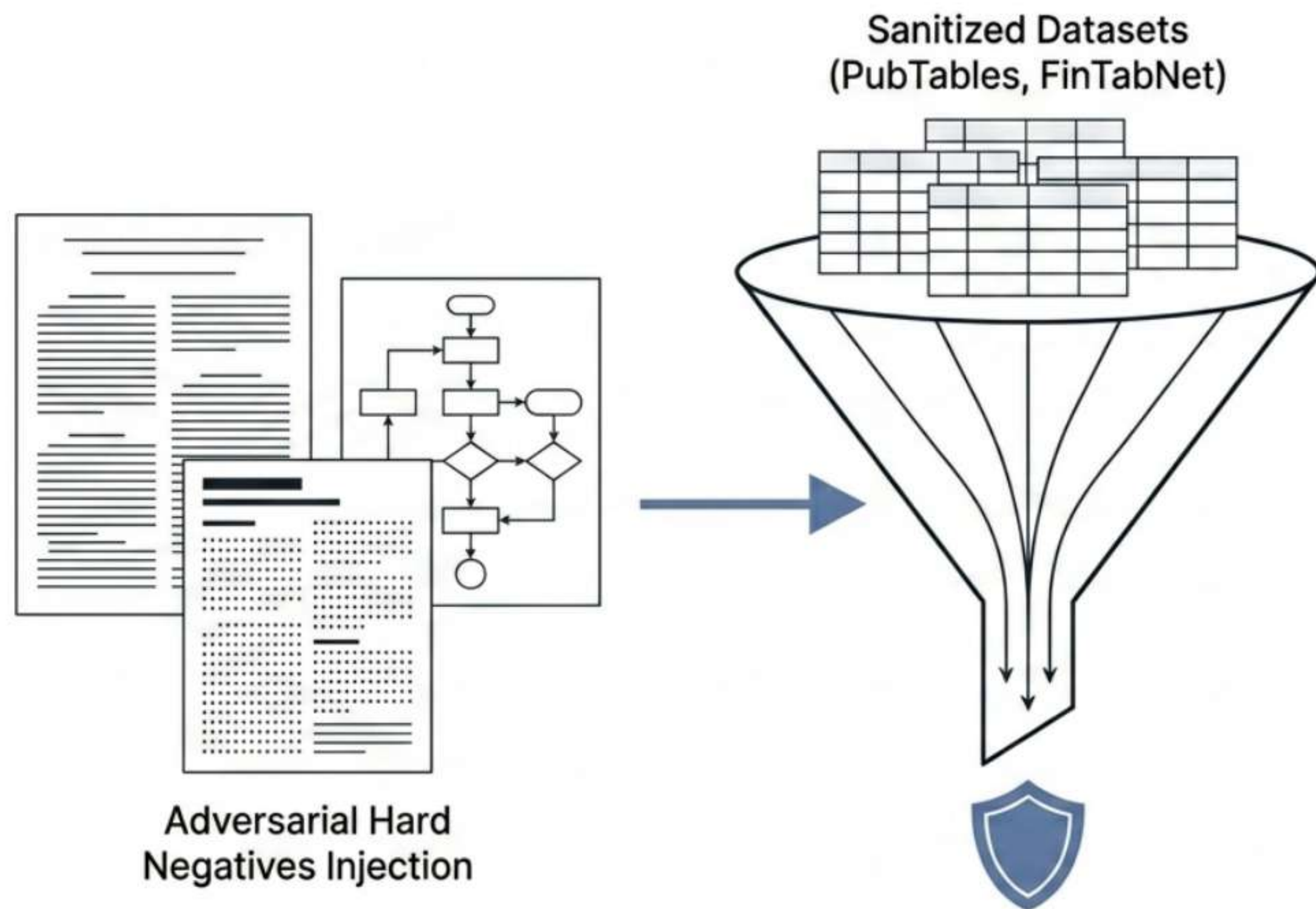


Moving to DETR architectures solved the semantic gap. Models must recognize structural relationships and negative space, not just localized pixel patterns.

Secondary constraint: YOLO's GPL license forced an architectural pivot to maintain proprietary deployment flexibility.

The Synthetic Dataset Odyssey

Training for reality via adversarial engineering



Competitors train on sanitized data, creating yes-men models that hallucinate tables on text paragraphs.

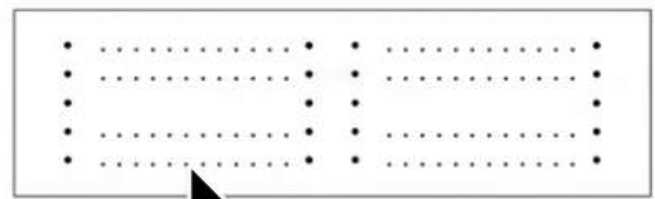
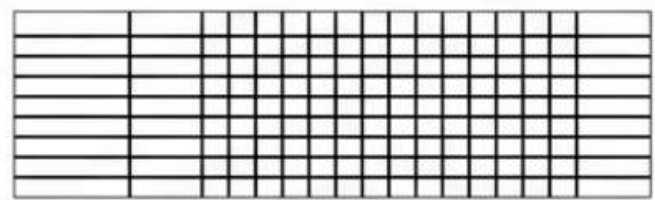
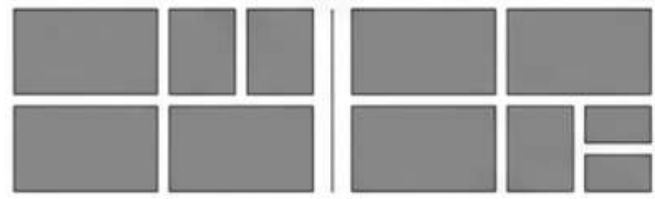
We inject custom-generated adversarial noise (hard negatives) directly into the pipeline.

The model learns to aggressively reject non-table structures, drastically lowering false-positive rates in production.

The Synthetic Dataset Odyssey

High-quality datasets were too beautiful.

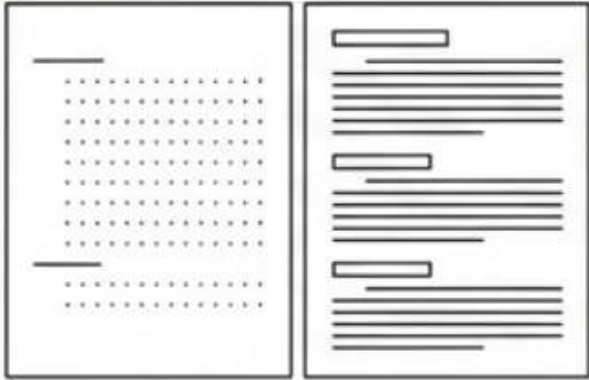
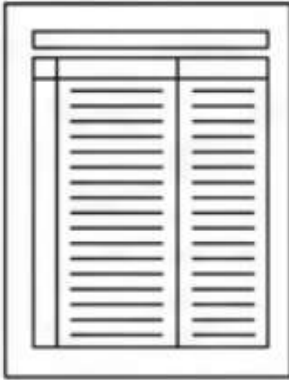
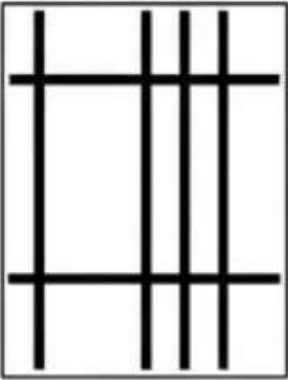
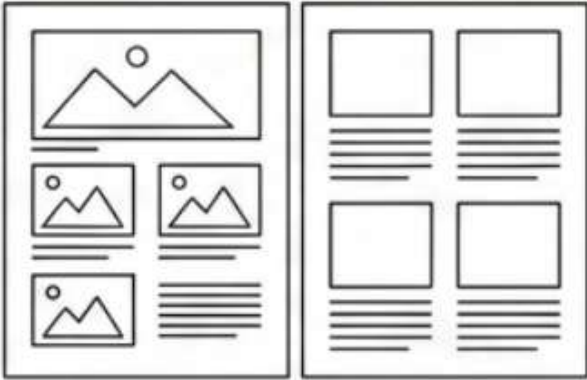
The model needed hard negatives “like images with intentionally empty label files” to teach the model the concept of pure background

12%	Tier 1: Sparse grid-like structures (e.g., Tables of Contents with dotfills).	
23%	Tier 2: 2- or 3-column structured lists mimicking tables.	
38%	Tier 3: Heavy use of vertical/horizontal rules without tabular semantics.	
27%	Tier 4: Grid-arranged sub-figures and side-by-side comparison layouts.	

The Synthetic Dataset Odyssey

Defeating the yes-man model via adversarial synthesis

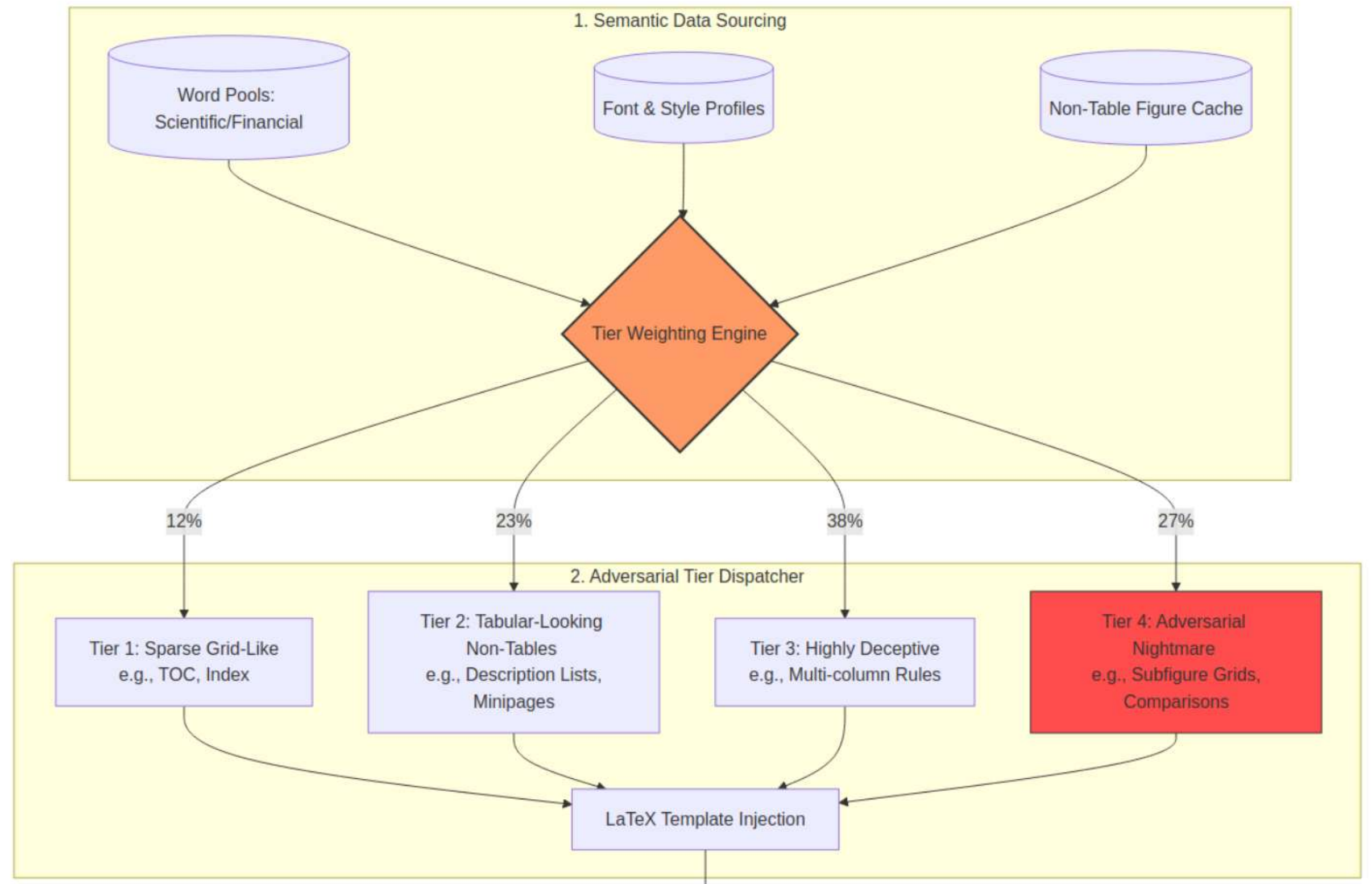
Metric	Sanitized-Only Training	Adversarial Mixed Training
Dataset Makeup	PubTables Only	Phase 3 + Synthetic
Benchmark AP	0.93	0.89
Real-World FPR	High	Low
Model Behavior	'Yes-Man' (over-predicts)	Discriminative (rejects noise)

Tier 1 (12%)	Tier 2 (23%)	Tier 3 (38%)	Tier 4 (27%)
			

The Synthetic Dataset Odyssey

Dataset Automation Engine

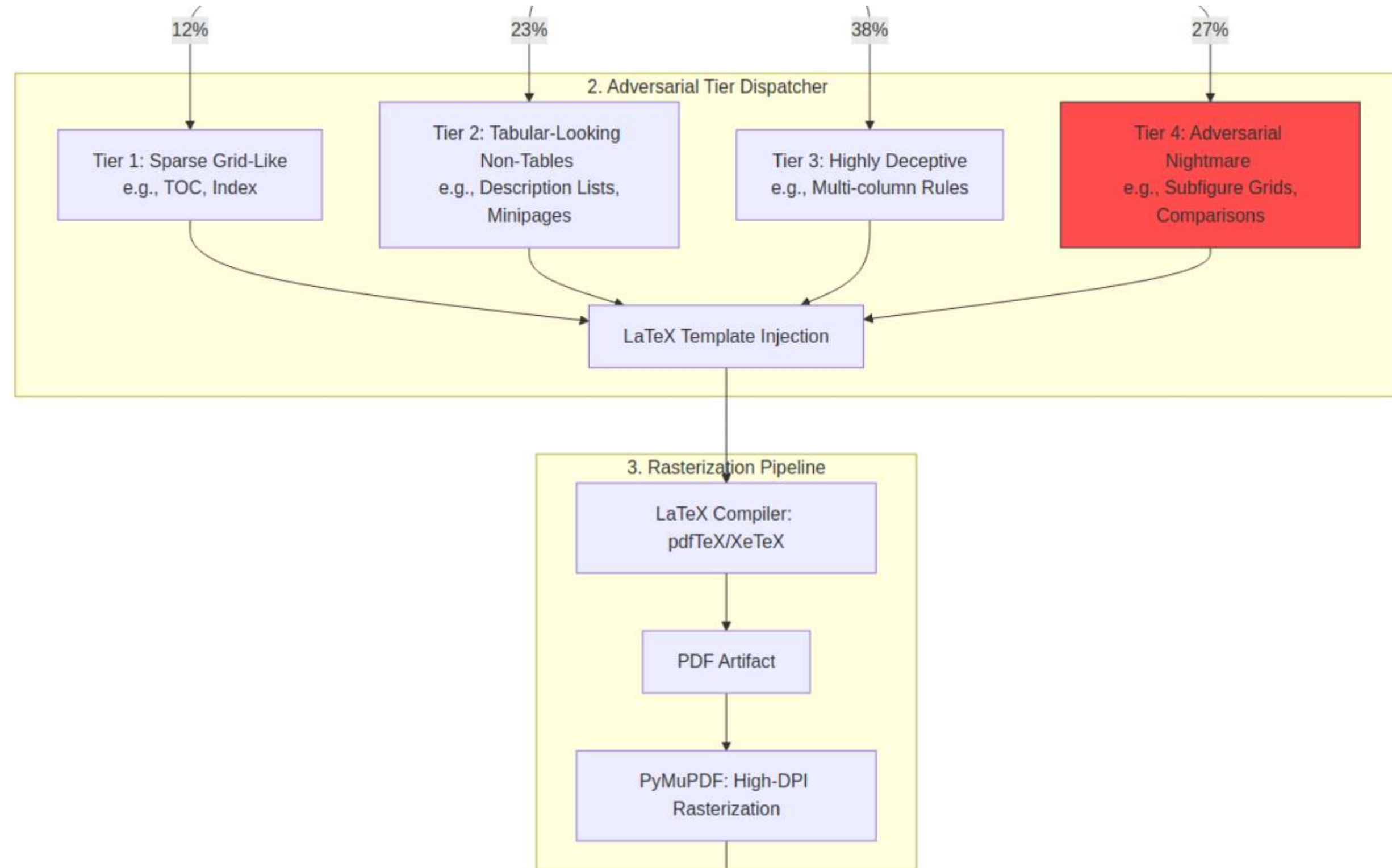
Part 1



The Synthetic Dataset Odyssey

Dataset Automation Engine

Part 2



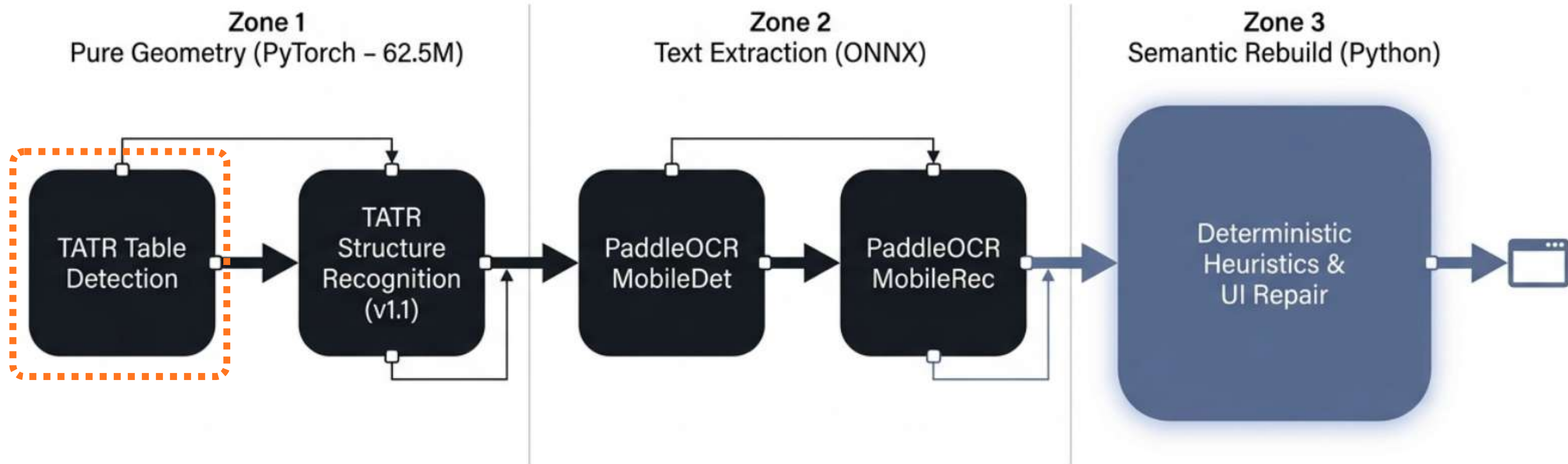


Table Detection

Table Detection

Core Philosophy: Robustness Over Sanitized Metrics

The primary objective was to build a production-ready table detector, not a leaderboard-overfitter.

Training solely on sanitized, easy datasets (PubTables/FinTabNet) yielded an artificially high AP > 0.93, but this model acted as a "yes-man" hallucinating bounding boxes on flowcharts and multi-column text due to the complete absence of hard negatives.

Table Detection

Core Philosophy: Robustness Over Sanitized Metrics

The fix was to inject the noisy, complex Phase 3 dataset (422 images) directly into the training and validation splits.

This dropped AP to ~ 0.89 , but fundamentally shifted the model from a naive bounding-box regressor into a highly discriminative system capable of rejecting false positive grids

Sanitized Data Only

The Trap: Training on pure PubTables yielded an artificially high AP > 0.93.

Result: Created a “yes-man” that hallucinated bounding boxes on flowcharts.

AP: 0.93+

Production Ready

Global Mixing

The Fix: Injecting 422 noisy Phase 3 images via a custom WeightedRandomSampler directly into the training split.

The Result: Intentionally dropped theoretical AP, shifting the model into a highly discriminative system capable of rejecting false positives in the wild.

AP: ~ 0.89

AP50: 0.975

Table Detection

Data Engineering: Global Mixing & Weighted Sampling

Sequential training (Dataset $A \rightarrow B \rightarrow C$) was actively avoided. Neural networks suffer from catastrophic forgetting; fine-tuning a converged model exclusively on 422 noisy images would overwrite foundational knowledge.

Table Detection

Data Engineering: Global Mixing & Weighted Sampling

Global Mixing: Merged massive clean datasets (Phase 1, cTDaR) with the small noisy Phase 3 dataset into a single ~4,000+ image distribution.

Weighted Oversampling: A custom WeightedRandomSampler forced the dataloader to dynamically oversample the minority dataset (Phase 3 + hard negatives) in every batch forcing the network to satisfy both simple grids and complex edge cases simultaneously.

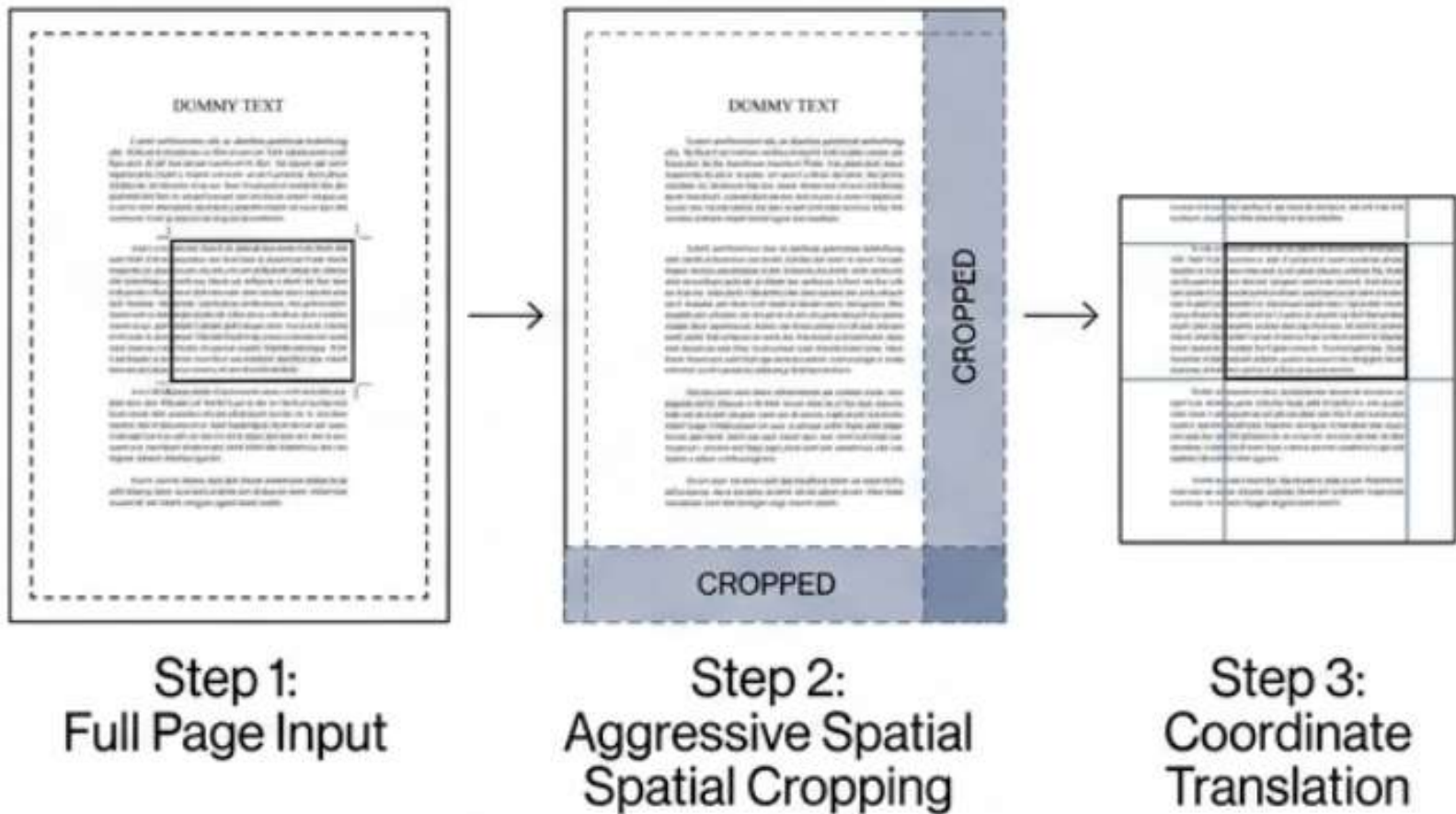
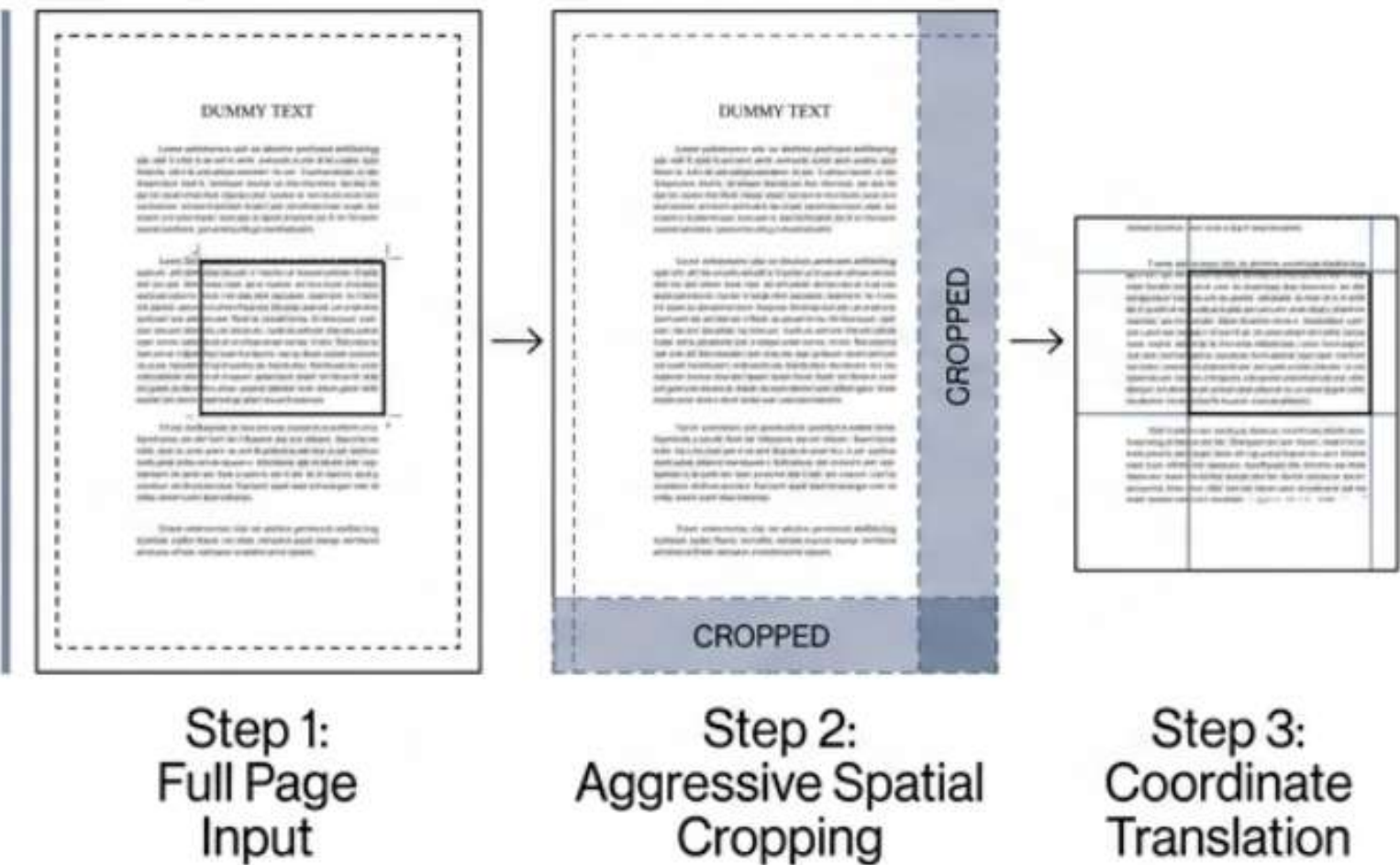
Table Detection

Spatial Regularization vs. Positional Memorization

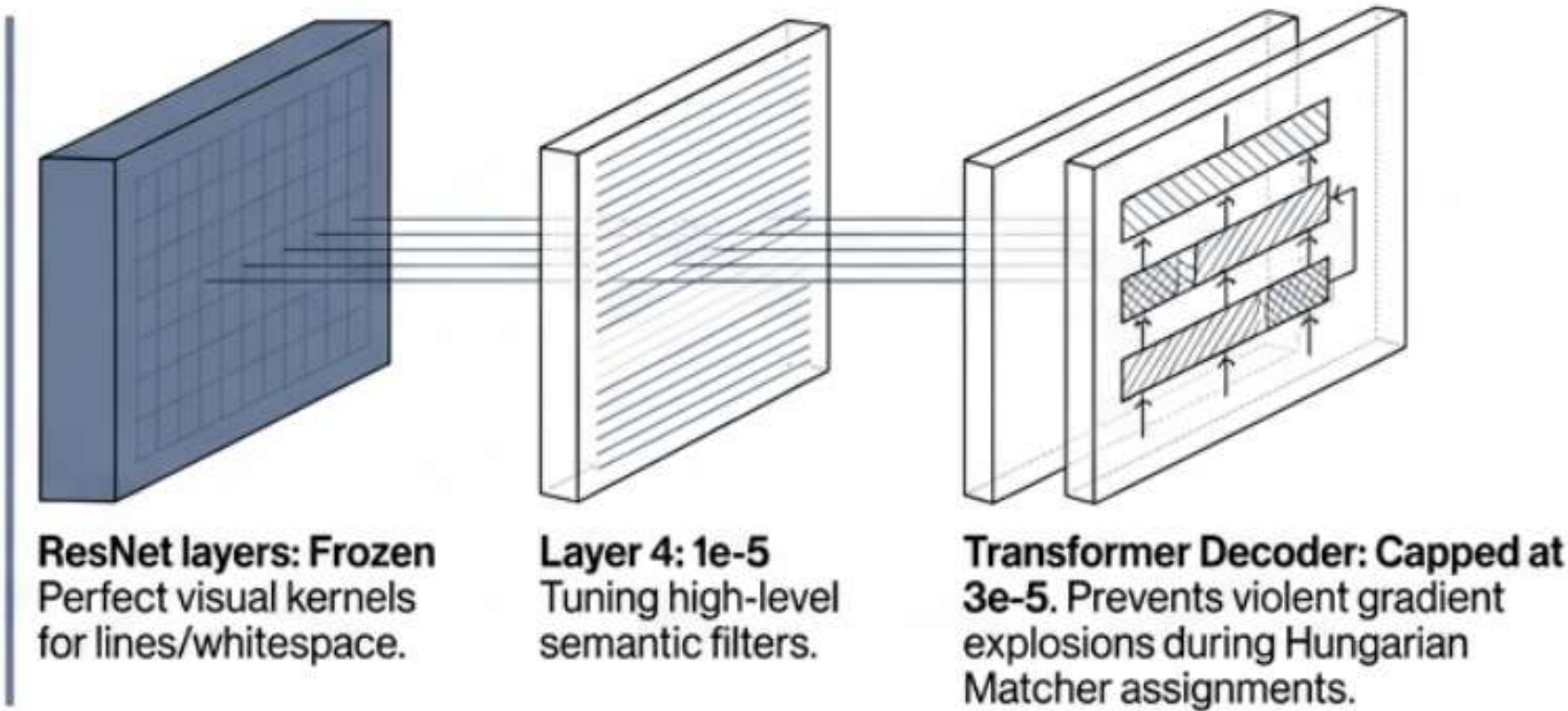
DETR architectures rely heavily on absolute positional embeddings. Trained solely on full-page images, the model learns to predict from predictable page margins and fails catastrophically on tightly cropped documents.

Forcing translation invariance and surgical backbone freezing

Spatial Regularization



Differential Learning Rates



DETR architectures rely on absolute positional embeddings. Random Spatial Cropping (up to 15% margin removal) forces true translation invariance over page-edge memorization.

Table Detection

TATR TD training "Datasets used: mixed "pubtables + tablebank word + ctar + hard negatives synthetic

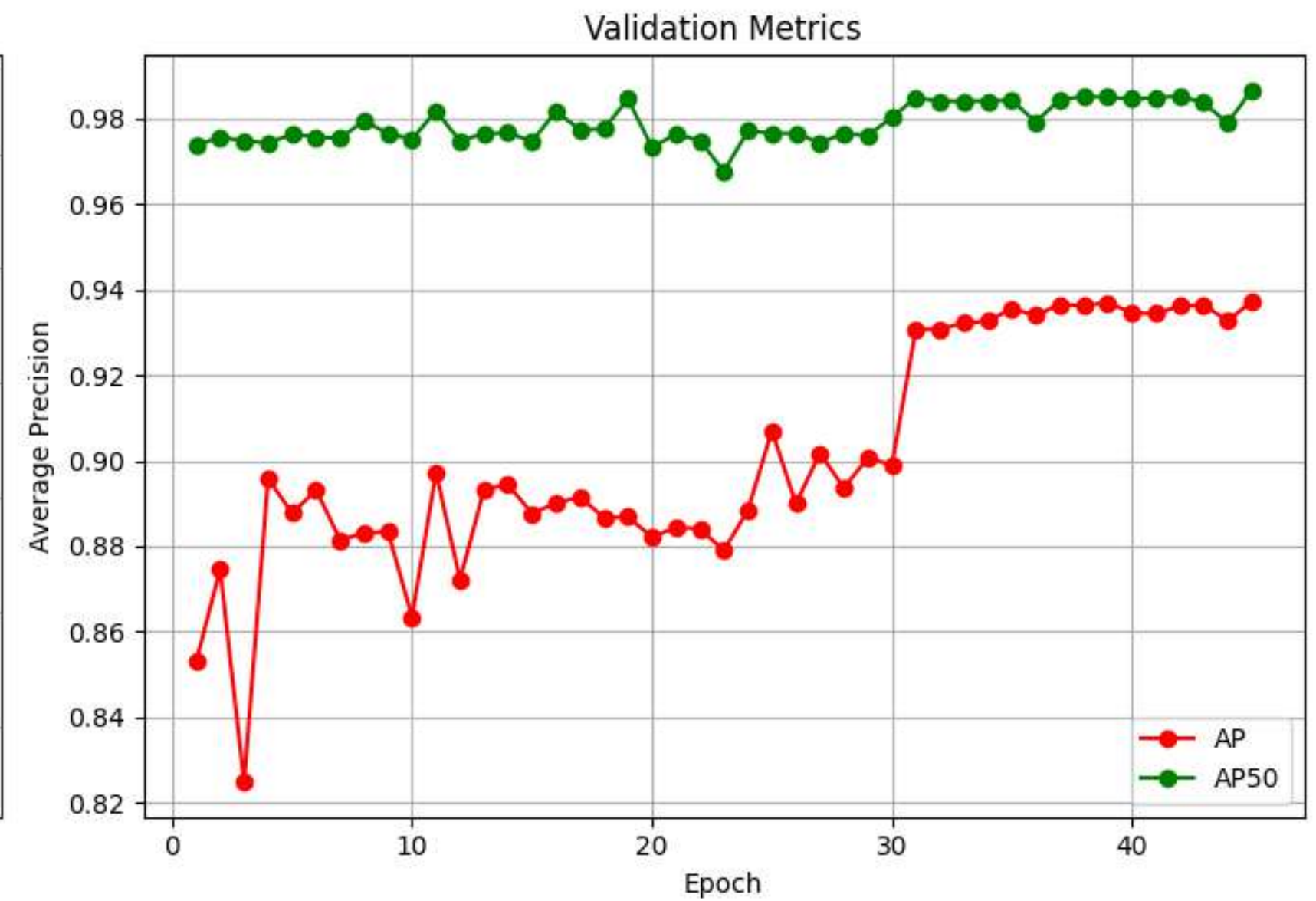
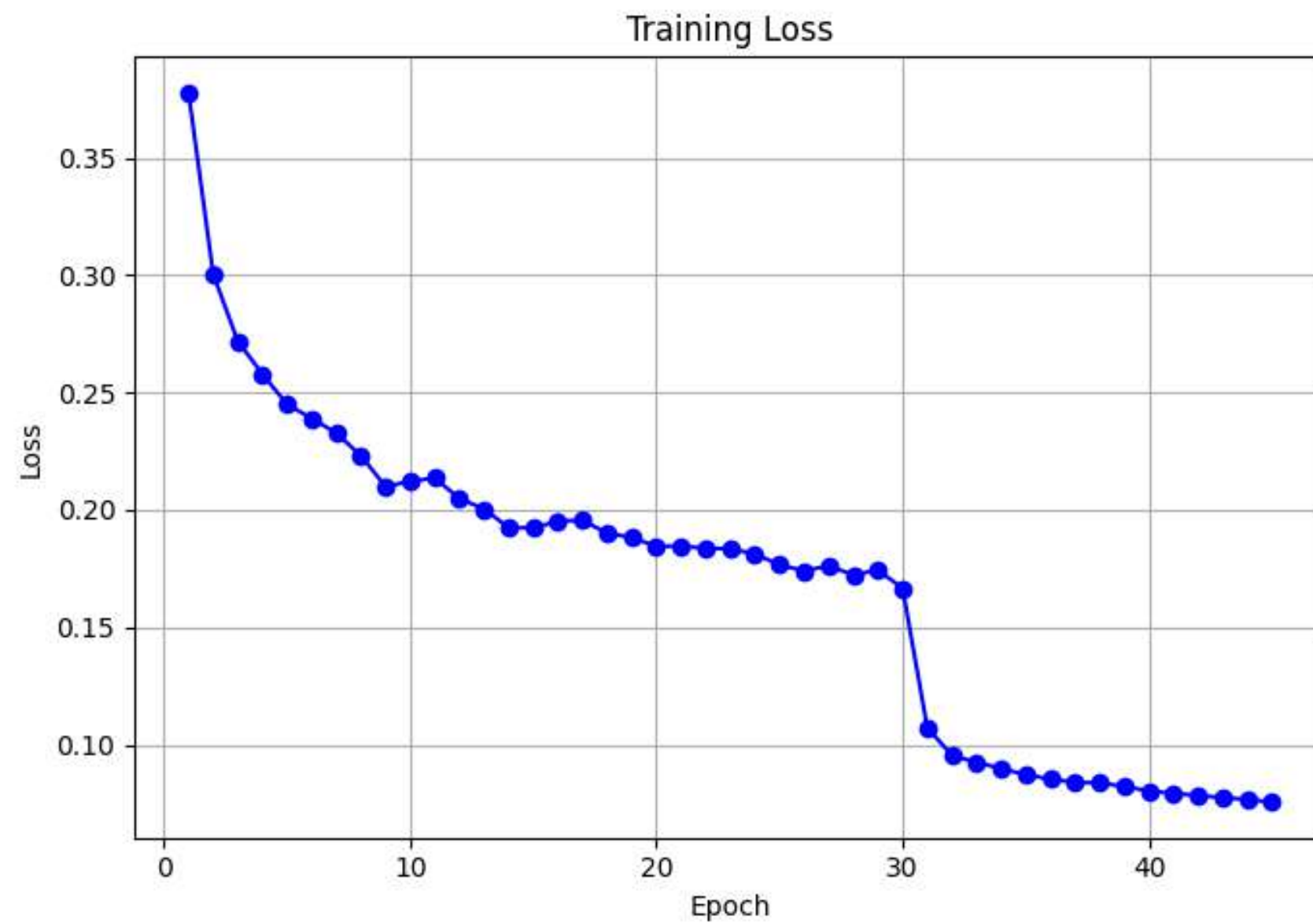


Table Detection

Differential Learning Rates & Surgical Freezing

microsoft/table-transformer-detection is a heavily converged foundation model. Treating it as a from-scratch training run destroys pre-trained kernels.

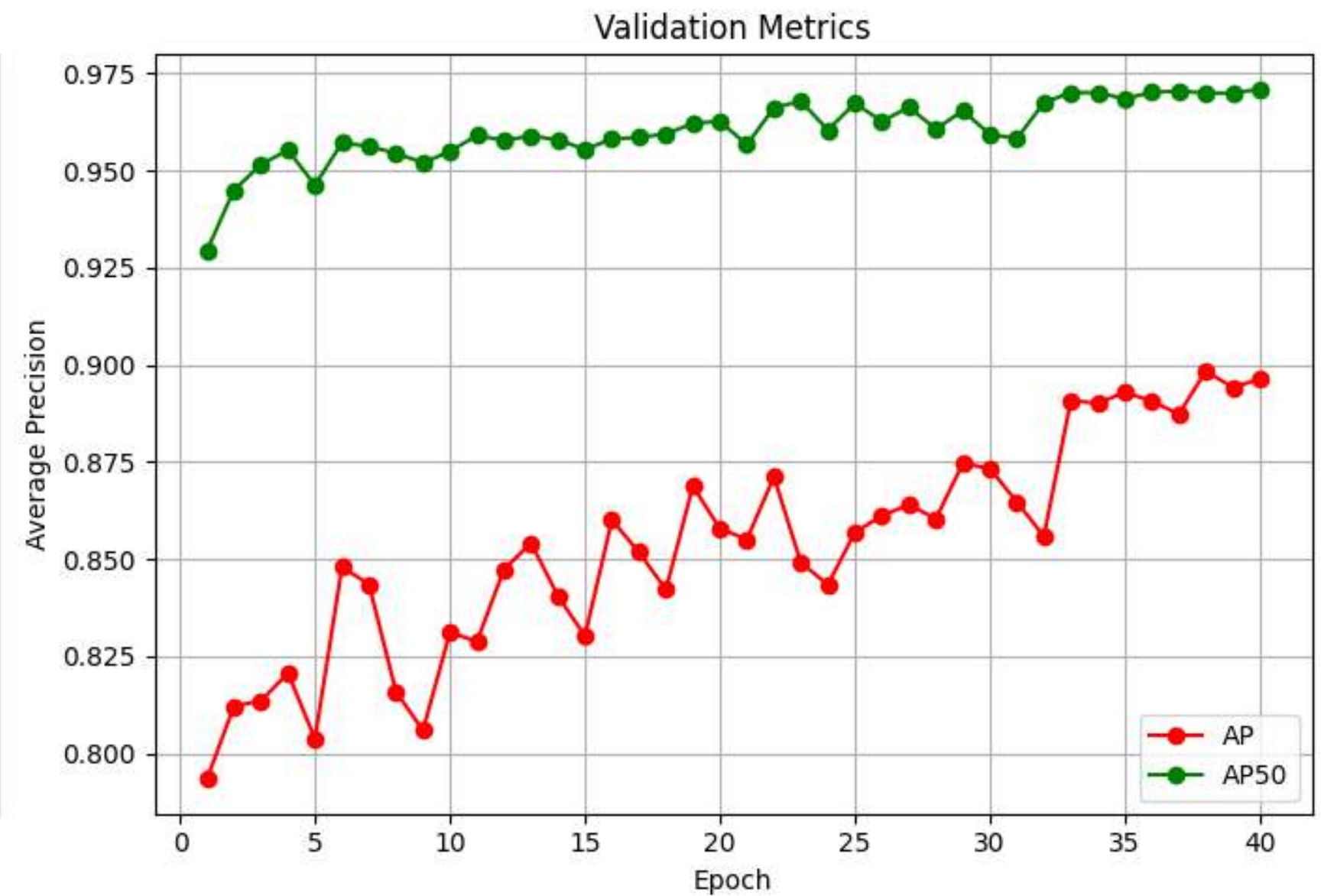
Backbone Freezing: Lower-level ResNet layers were frozen, they already possess perfect visual kernels for lines, whitespace, and text textures.

Top-Layer Tuning: layer4 was unfrozen with a differential learning rate ($LR_BACKBONE = 1e-5$). With ~4,000 images, sufficient data volume existed to safely update high-level semantic filters without gradient shattering.

Transformer Restraint: The Transformer Decoder learning rate was capped at $3e-5$ (resisting the standard $1e-4$), preventing violent gradient explosions during the Hungarian Matcher's bipartite assignments.

Table Detection

TATR TD training "Datasets used: mixed "pubtables + tablebank word + ctidar + hard negatives synthetic + Phase3 dataset



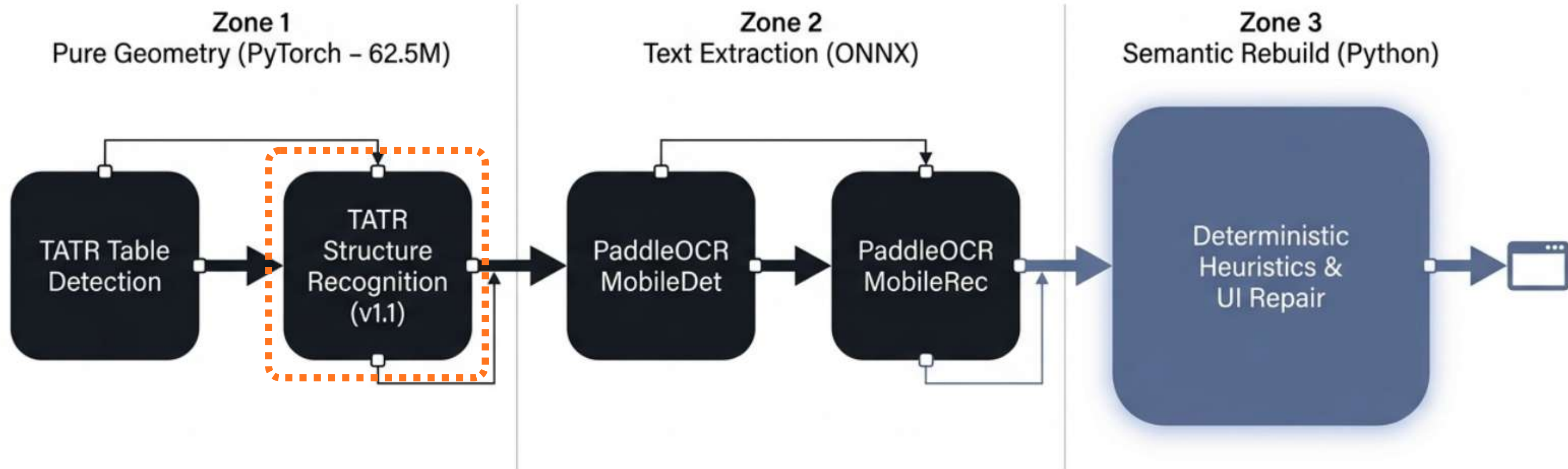


Table Structure Recognition

Table Structure Recognition

The TSR journey was the most turbulent part of the project:

SLANet / SLANeXt — tried all variants after painful annotation format conversions.
Didn't impress enough.

TATR TSR v1.1 — emerged as the best fit. Surfed research papers on UniTable, TableFormer, LORE, Nougat, Donut, Florence2, but TATR struck the best balance between parameter count and performance.

Table Structure Recognition: The Phase 3 Annotation Disaster

EVIDENCE CROP 1: FINANCIAL TABLE

[illegible]

EVIDENCE CROP 2: LEASE TABLES

Financing Arrangements		Requirements	
Year Ended	Principal	Interest	
August 31,			
2026	\$ 775,996	\$ 151,132	
2027	811,231	115,545	
2028	848,143	78,837	
2029	886,733	40,346	
Totals	\$ 3,322,022	\$ 386,263	\$ 3,708,315

Financing Arrangements include leases payable and a summary of term leases payable for the year ended August 31, 2025, is as follows:

Issue	Description	Rate	Value	Date	Payable
8/1/2023	Printers	1.649%	412,615	2028	251,609
					\$ 837,577

In the event of default, the outstanding lease payable is secured by the leased assets. The future payment requirements for the leases payable are as follows:

The Trap

- Initial model evaluation yielded a disastrous AP of **0.477**.
- Gradients exploding across the decoder layer.

The Audit

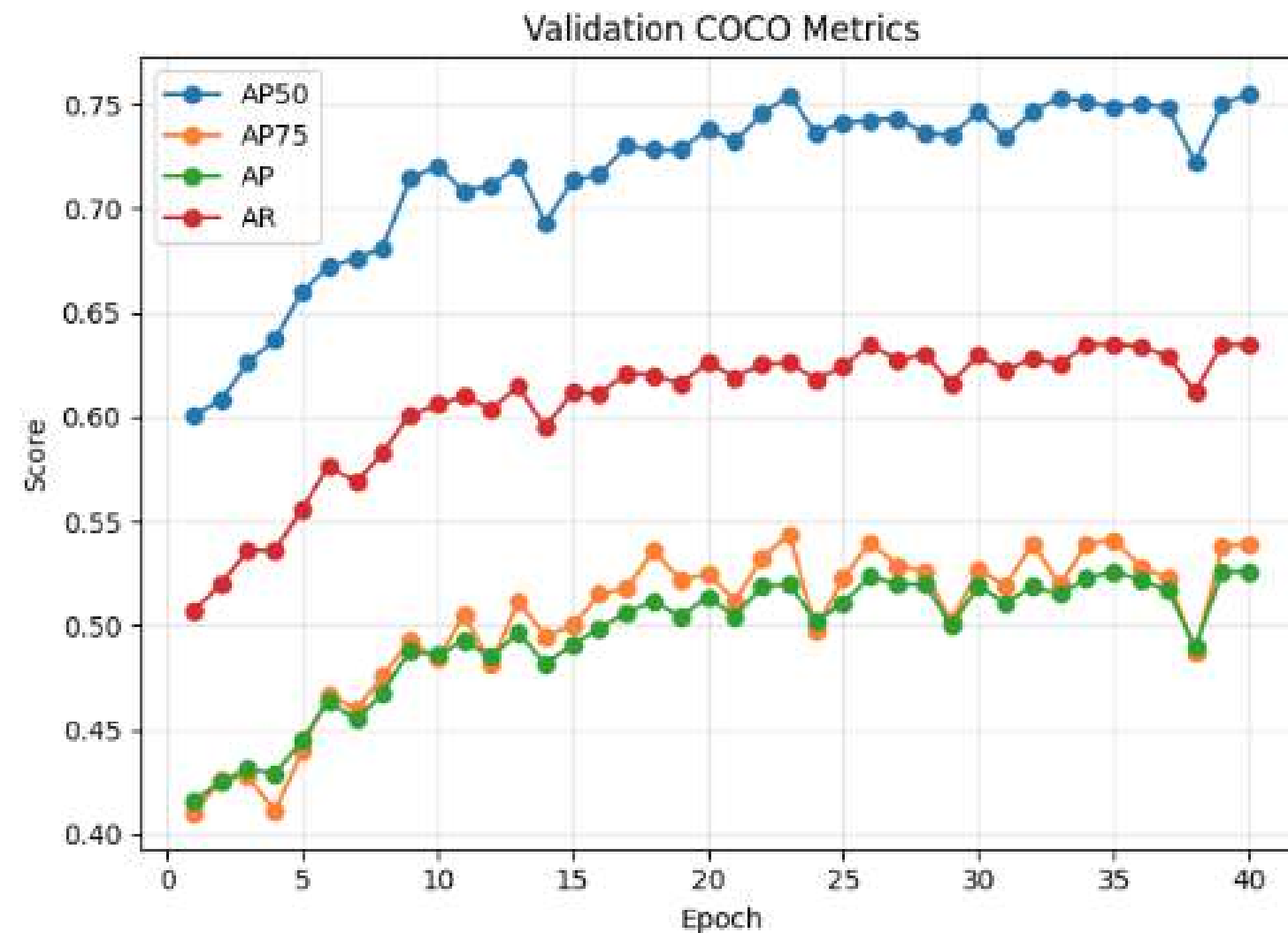
- Over 60 images contained **5,000+** hallucinated projected row header annotations in overlapping chaos.
- Column headers entirely missing. Spanning cells heavily misused.

The Ablation Proof

Temporarily dropping the class caused an instant ~20% AP jump, proving the dataset was mathematically poisoning the latent space.

Table Structure Recognition: The Phase 3 Annotation Disaster

Before discovering the annotation issue



model: TATR TSR v1.1 all.

Datasets used: 3000 pubtables 1M, 4000 Fintabnet.c, 422 phase3 dataset"

The Trap

- Initial model evaluation yielded a disastrous AP of **0.477**.
- Gradients exploding across the decoder layer.

The Audit

- Over 60 images contained **5,000+** hallucinated projected row header annotations in overlapping chaos.
- Column headers entirely missing. Spanning cells heavily misused.

The Ablation Proof

Temporarily dropping the class caused an instant **~20% AP jump**, proving the dataset was mathematically poisoning the latent space.

Table Structure Recognition: The Phase 3 Annotation Disaster

One-Click Fix:

Using AI-generated annotation app

- Deleted **5353** false annotations

Bulk Delete

Category

table projected row header ▼

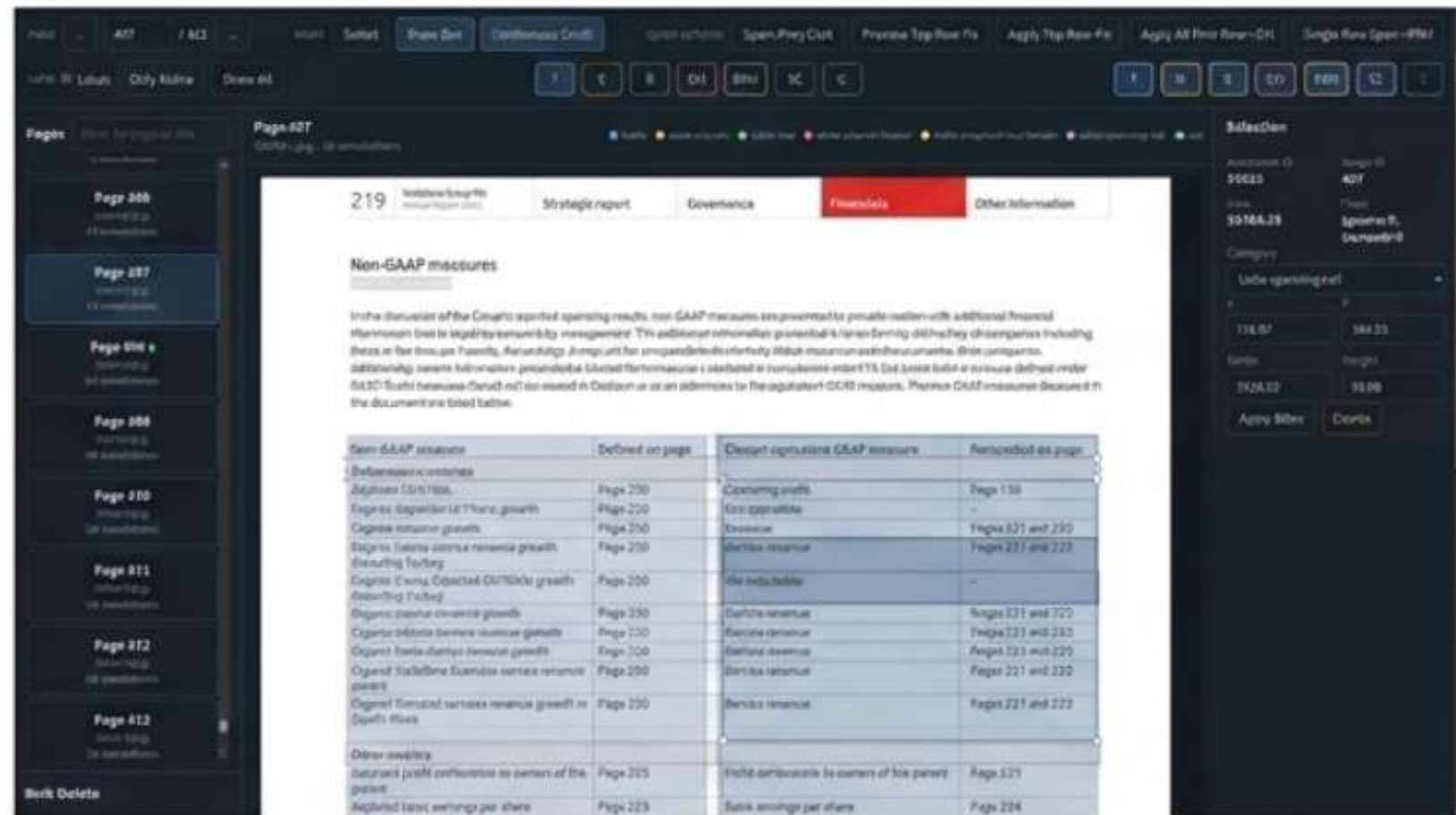
Start Page End Page

82 154

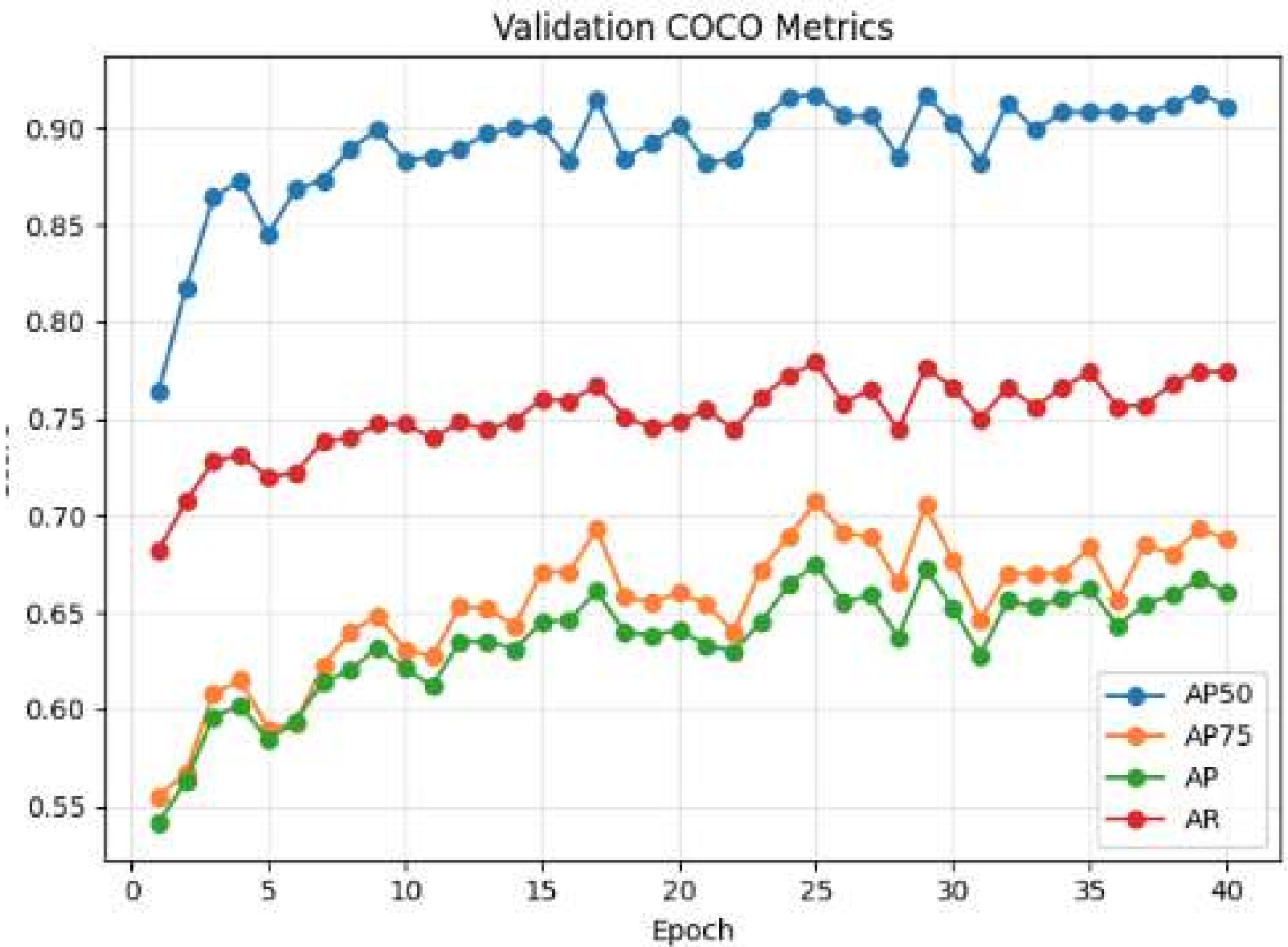
Preview Delete Range

Deleted 5353 annotations from pages 82-154.

Tool-assisted ground truth overhaul

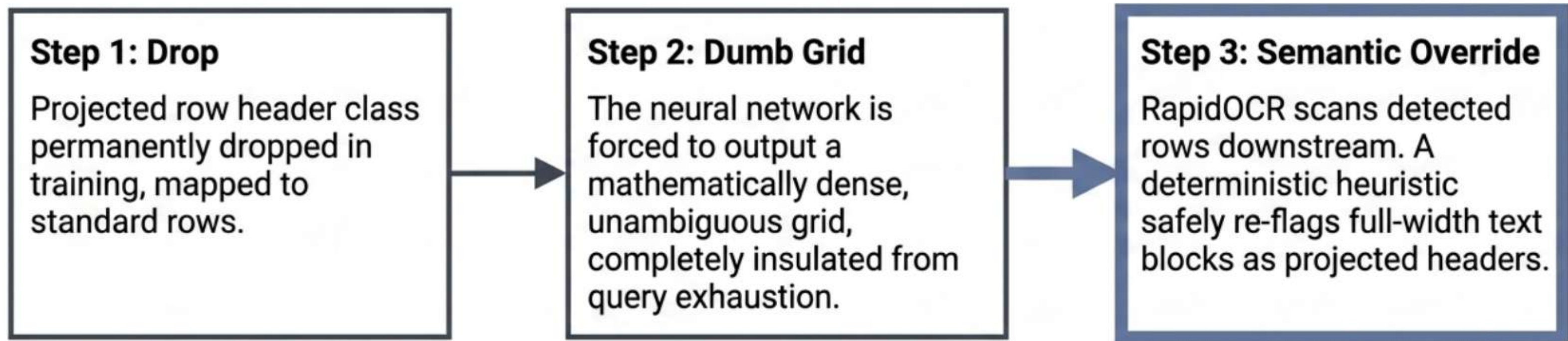


The Intervention: Building an AI-assisted editor for surgical repair. **5,000+** toxic annotations nuked in a single click. Programmable heuristics retroactively repaired spanning logic.



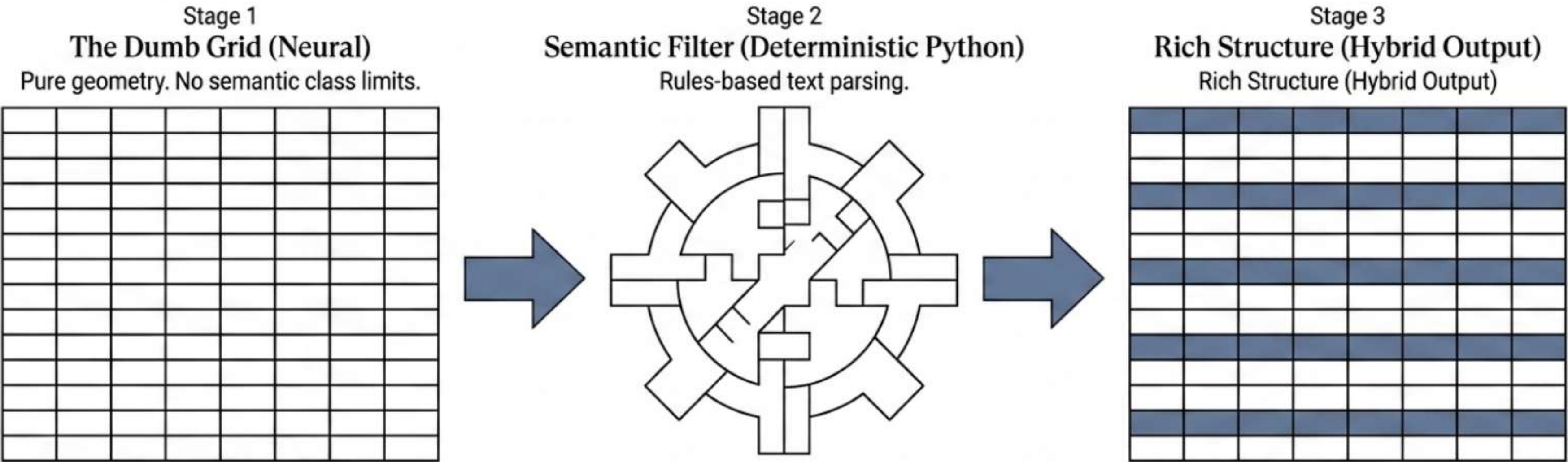
Results after dropping projected row header from classes

Table Structure Recognition: The 125-Query Algorithmic Bypass



The Bottleneck: TATR's decoder is strictly bound by a 125 `num_queries` limit. Intersecting phantom grids from row headers devoured this budget on large tables.

Offloading semantics to deterministic heuristic layers



The Pivot

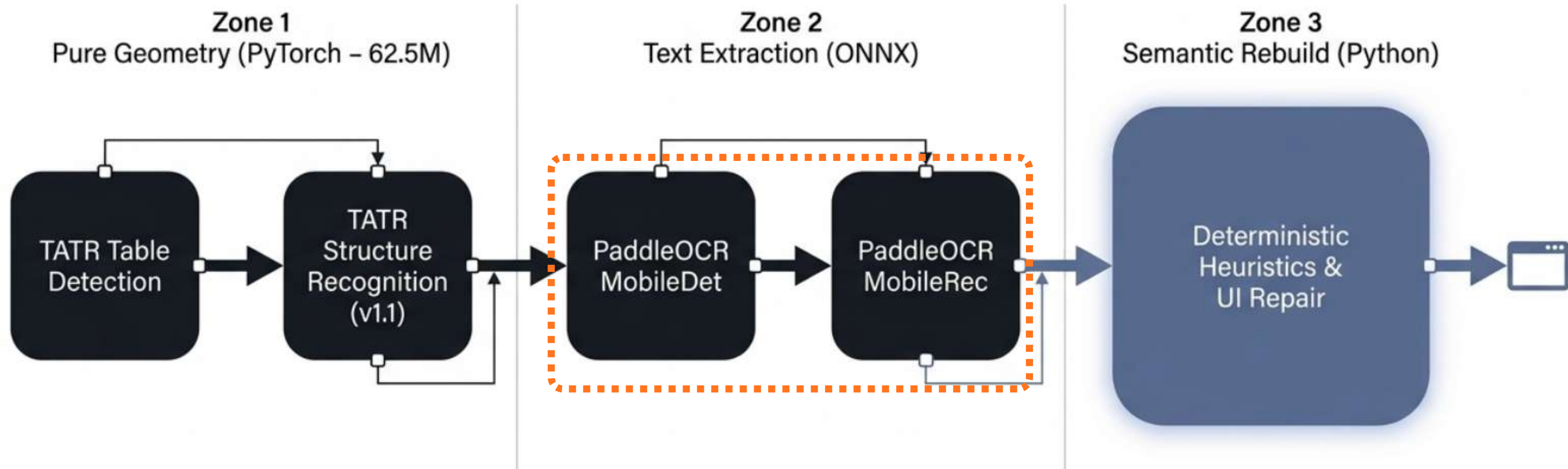
The projected row header class was intentionally dropped permanently. All annotations were mapped to standard rows.

The Result

Forces the neural network to output a mathematically dense, unambiguous grid, completely insulating it against query exhaustion.

The Semantic Heuristic

Downstream RapidOCR parses the rows. If a row contains a single continuous text block spanning the table width, deterministic code safely re-flags it as a projected row header.



OCR & Heuristics Integration

OCR & Heuristics Integration: Not Reinventing the Wheel

After inspecting the full PaddlePaddle Document AI pipeline, **ppocrv5_mobile_det** and **ppocrv5_mobile_rec** were selected.

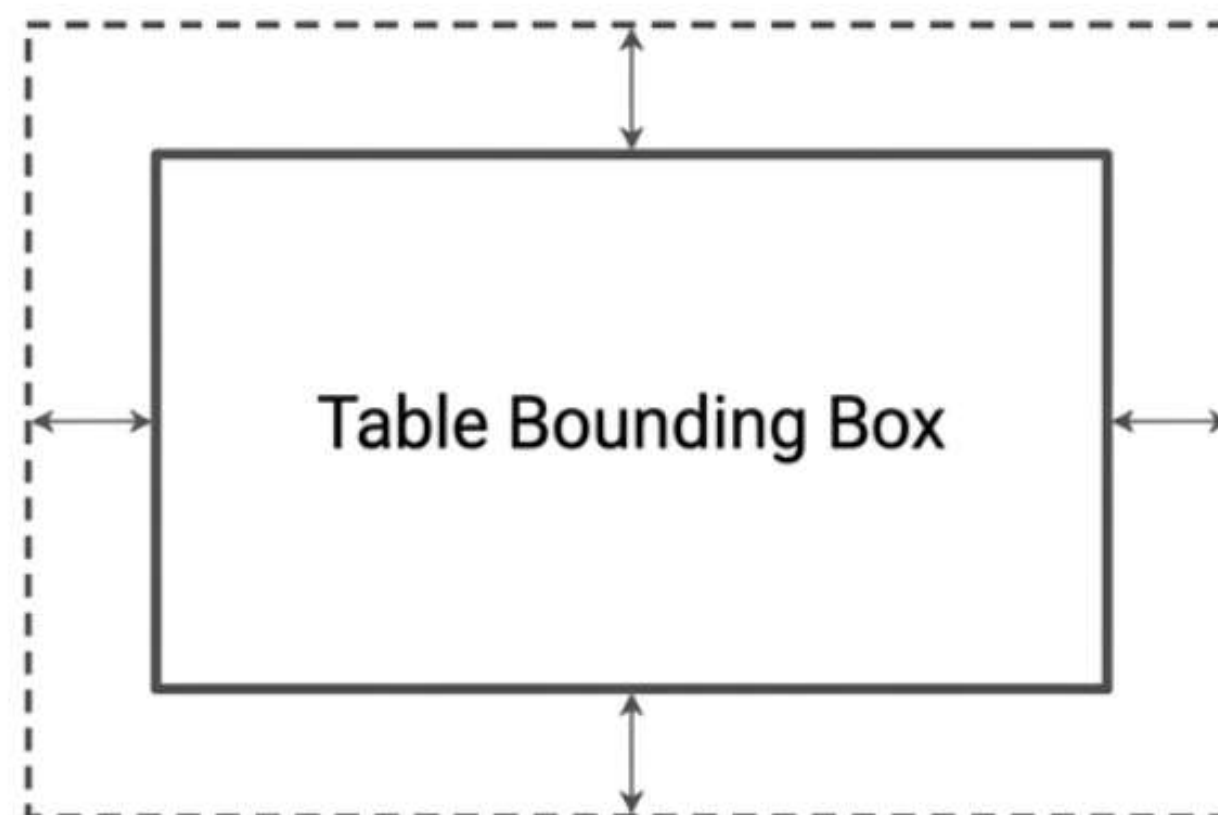
Retraining was unnecessary given their quality and speed.

The det+rec separation also **unlocked specific UI/UX goals** for the studio editing experience.

OCR & Heuristics Integration

The Semantic Layer (RapidOCR ONNX)

Engine Choice: PaddleOCR
v5 Mobile Det+Rec was
astonishingly fast;
Retraining was unnecessary.



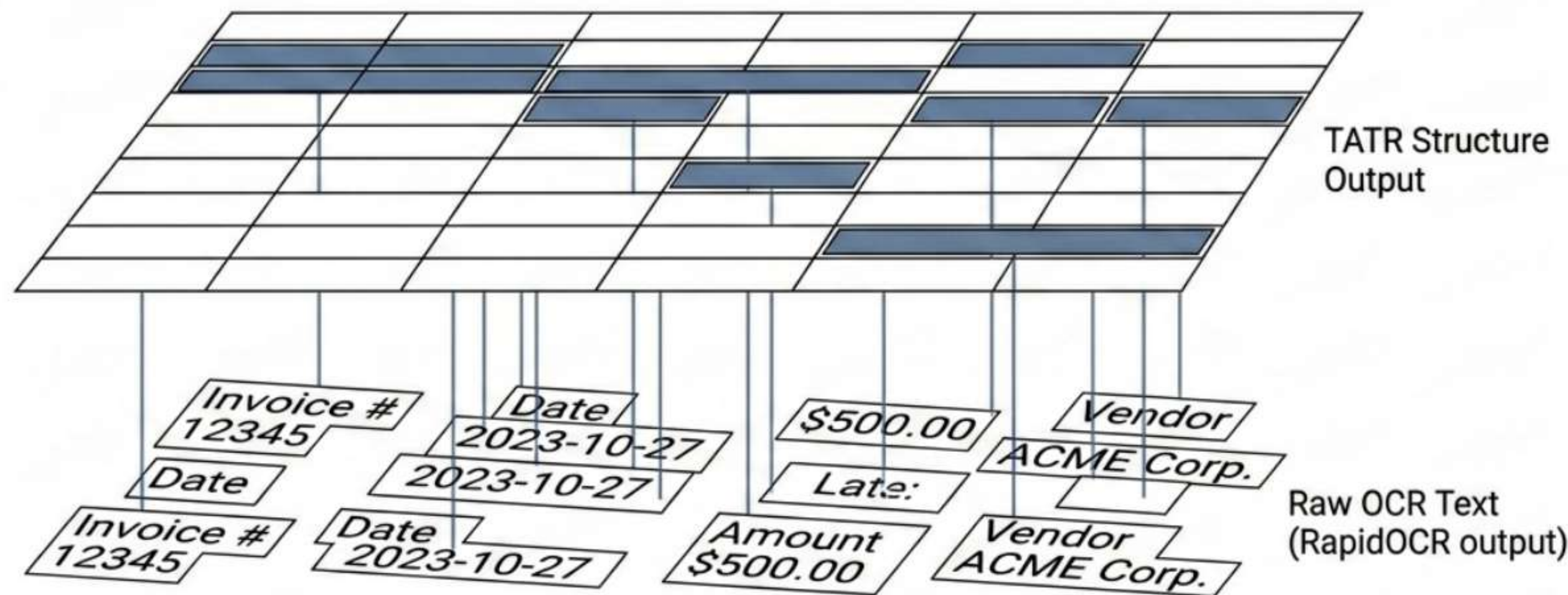
The Padding Heuristic
To prevent runaway
expansion into nearby
non-table text while giving
TSR context:

Base padding: 4%
Cross-axis cap: 20%
Clamp: min 2px, max 24px

Assignment: OCR items assigned to cells via robust scoring:

$$\text{score} = \text{ioa} + 0.20 * \text{iou} + 0.35 * \text{center_inside}$$

OCR & Heuristics Integration: Cell Assignment Logic

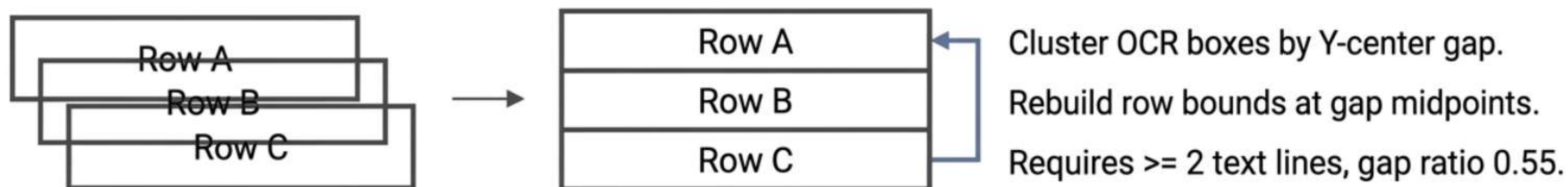


Primary Scoring Equation	Acceptance Thresholds	Fallback Mechanism	Refinement
$\text{score} = \text{ioa} + (0.20 * \text{iou}) + (0.35 * \text{center_inside})$	$\text{IoA} \geq 0.35 \text{ OR } \text{IoU} \geq 0.02$	Nearest-cell-center assignment (max normalized distance 0.60).	Text-anchored row rebuild re-aligns bad TSR row boundaries by anchoring them to OCR text line Y-center gaps.

Deterministic Rebuild Logic

Problem: TSR models occasionally output misaligned row boundaries.

Solution 1: Text-Anchored Row Rebuild



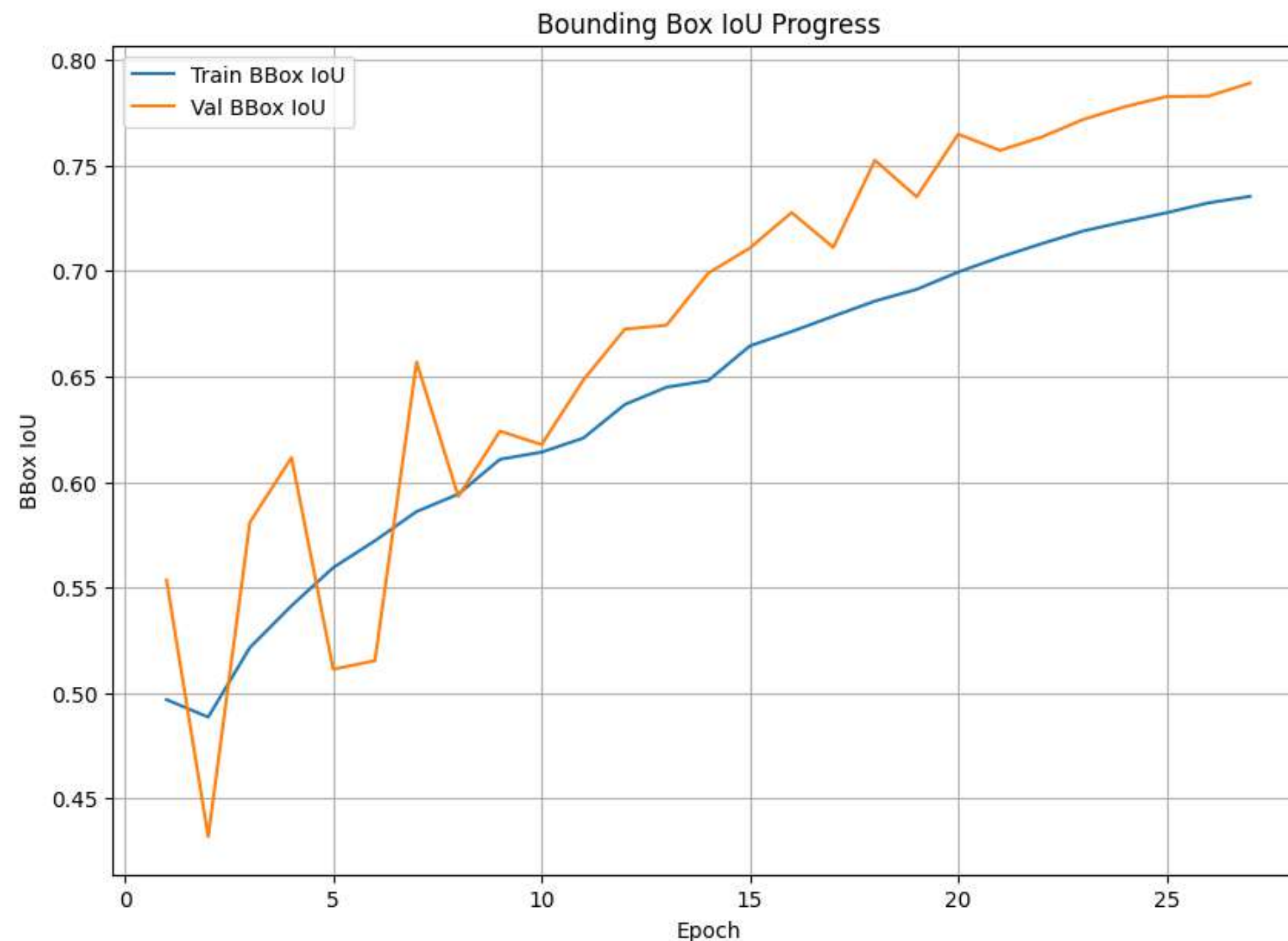
Solution 2: Left-Anchored Spanner Merge



Honorable Mention

Honorable Mention

Evaluating Alternatives: IBM's Docling TableFormerV2



Out-of-box Excellence

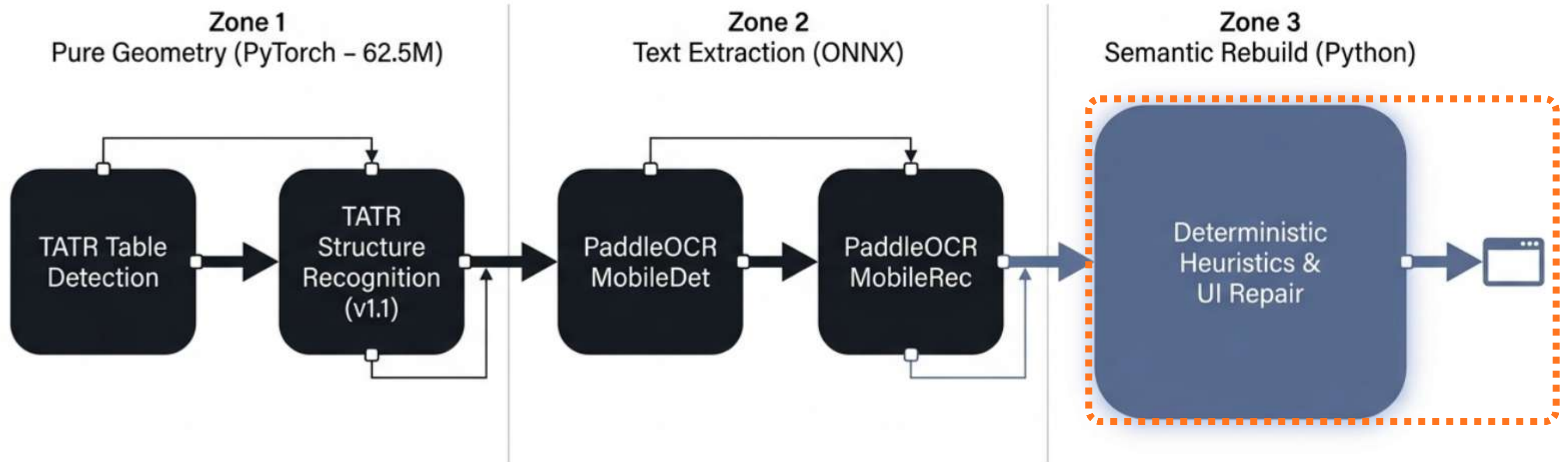
52M parameter model produced flawless results on standard tasks without any training.

The Limitation

The training pipeline fundamentally expected clean, text-annotated bounding boxes. It proved incompatible with the noisy, chaotic nature of the Phase 3 dataset.

The Decision

Retained strictly as a wrapped auxiliary model with custom post-processing, while the 62.5M TATR pipeline remained the primary engineered solution.



The Product Experience & Studio

The surgical table extraction studio.

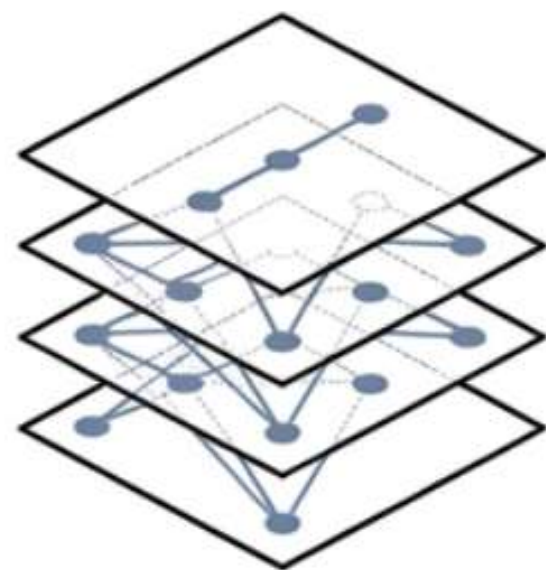
Agent P-DF: Enterprise document extraction built for the final mile.



Agent P-DF

“AI is imperfect, but the editing environment should be perfect.”

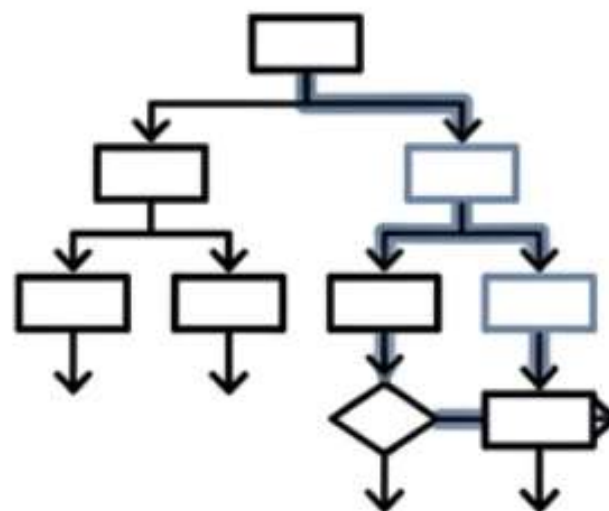
Pillar 1: Neural Networks



62.5M total parameters.
Handles pure geometry.

Operates entirely locally within the
lightweight gap left by monolithic 900M+
parameter models like PaddleOCR VL 1.5.

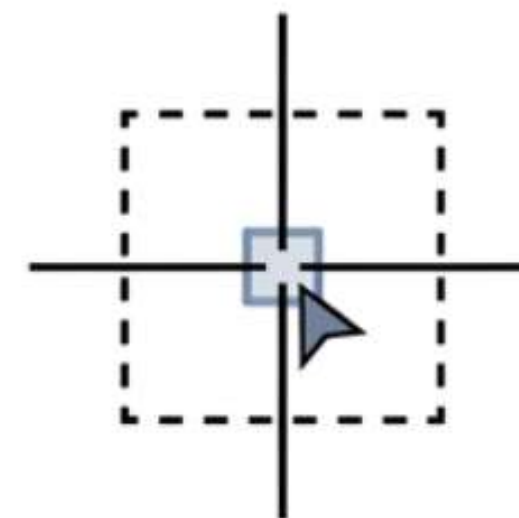
Pillar 2: Deterministic Heuristics



Python-based rule sets handle
semantic logic.

Programmatic guardrails that catch and
correct algorithmic edge cases
downstream.

Pillar 3: The UI / Human

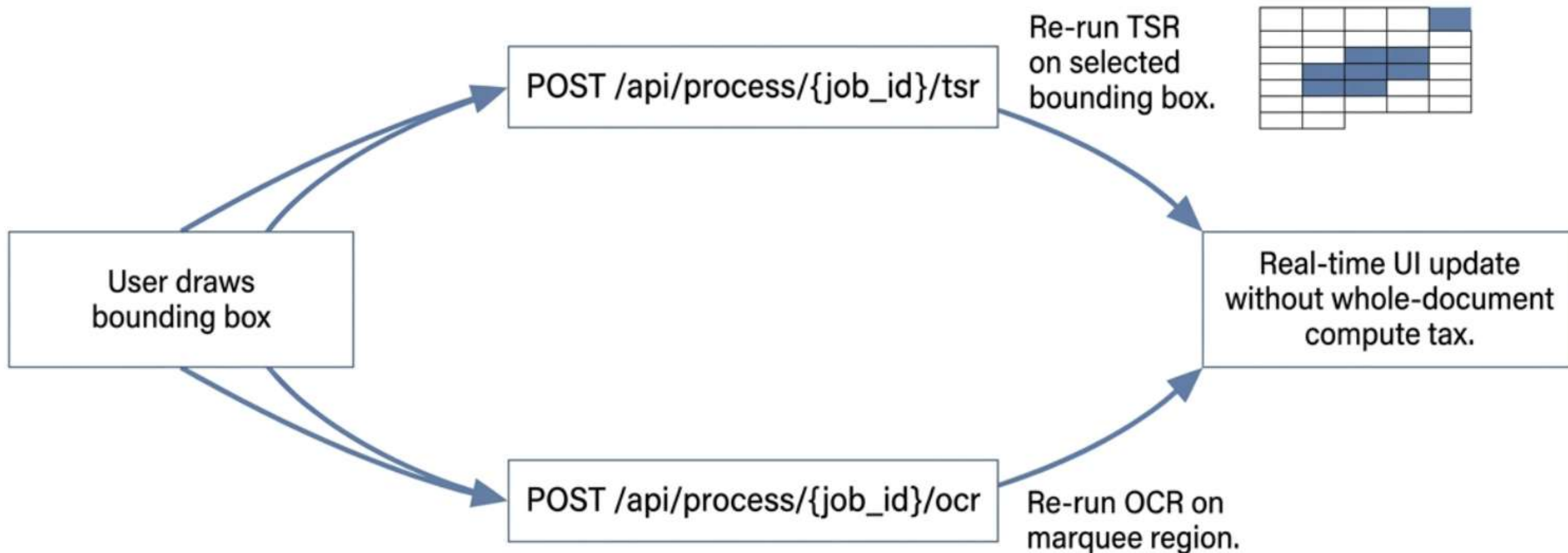


The interactive editing environment
handling the final mile.

Surgical toolsets for localized reruns
and structural mutation.

Localized Reruns: The Correction Loop

The system architecture empowers the UI to be highly targeted. Instead of forcing costly full-document re-runs when an error is found, users can trigger localized API calls that surgically reprocess specific bounding boxes on demand.



Landing Page



Agent P-DF

[Get Started](#)



Agent P-DF

Extract Data with **Precision**

Unleash the supreme AI table extractor. Upload complex PDFs, and watch our intelligent agent seamlessly reconstruct them into flawless, usable tables.

[Try Table Detective Free](#)

Home Page

Extract Tables with AI Precision

Upload document images or PDFs and let our 5-phase pipeline detect tables, recognize structure, perform OCR, and extract editable data — all in seconds.



Drop files here or click to browse

Supports JPG, PNG, BMP, TIFF, and PDF up to 30 MB each

0 accepted

0 rejected

0 active

1

Table
Detection

2

Structure
Recognition

3

Text Detection

4

OCR
Recognition

5

Cell
Assignment

User Guide

1. Upload files

Drag and drop or browse to select one or many images/PDFs.

2. Wait for processing

Track queued and running jobs in the Upload Queue panel.

3. Review results

Open any completed job to edit table structure and cell text.

4. Export data

Export cleaned tables to CSV, Excel, or HTML from the studio.

The Editing Environment

Agent P-DF 000037.jpg

PAGE 1

Tool Box

Pages: 1/1 Speed: 18.2 sec/page Errors: 0

Back / New Upload

Export CSV

Export Excel

Export HTML

Viewing: Page 1

Table Table 1



DONALD J WILSON II | Account # 3251 7230 2544 | December 14, 2023 to January 16, 2024

Withdrawals and other subtractions - continued

Table 1		
ATM and debit card subtractions - continued		
Date	Description	Amount
01/08/24	MOBILE PURCHASE 0106 Reef Parking Miami FL	-40.00
01/08/24	CHECKCARD 0107 BB* UNDERDOGS CANTINA HTTPSBBOT.MENCA 24011344007000007761390	-74.56
01/08/24	MOBILE PURCHASE 0107 LAZ PARKING 680208-DP SAN FRANCISCOCA	-10.00
01/09/24	MOBILE PURCHASE 0107 DENNY'S #9622 LIVERMORE CA	-60.58
01/09/24	CHECKCARD 0108 APPLE.COM/BILL 866-712-7753 CA 24692164008108842503856 RECURRING	-9.99
01/09/24	MOBILE PURCHASE 0108 CHEVRON 0212470 TRACY CA	-30.52
01/10/24	CHECKCARD 0107 THE CLANCY, AUTOGRAPH SAN FRANCISCOCA 24692164009109564193065	-358.61
01/10/24	CHECKCARD 0110 CHEVRON 0212470 TRACY CA 24692164010109922776970	-46.02
01/12/24	Z VAPE N DISCO 01/12 #000002927 MOBILE PURCHASE 245 EAST 11TH STR TRACY CA	-25.00
01/12/24	FOOD MAXX #490 01/12 #000845946 PURCHASE FOOD MAXX #490 TRACY CA	-7.87
01/16/24	PMNT SENT 0113 CASH APP*DONALD WILSON* 800-9691940 CA 24793384013001491975084	-20.00
01/16/24	TRACY MARKET I 01/13 #000348400 MOBILE PURCHASE 15 E GRANT LINE R TRACY CA	-26.07
01/16/24	MOBILE PURCHASE 0113 MCDONALD'S F36229 TRACY CA	-25.18
01/16/24	PMNT SENT 0113 CASH APP*DONALD WILSON* 800-9691940 CA 24793384013002641296058	-25.00
01/16/24	PURCHASE 0114 ACE HARDWARE CORPORATIO 800-453-0660 IL	-17.30
01/16/24	\$CHECKCARD 0113 CHEVRON 0212470 TRACY CA 24692164013102943703625	-44.59
Total ATM and debit card subtractions		\$2,498.00

\$ There were not enough funds available in your account to cover the electronic transaction(s) received on 01/16/24. The transaction(s) indicated have been paid.

Table 2		
Other subtractions		
Date	Description	Amount
12/15/23	Zelle payment to MELISA Conf# zanwc9dna	-385.00



ATM and debit card subtractions - continued		
Date	Description	Amount
01/08/24	MOBILE PURCHASE 0106 Reef Parking Miami FL	-40.00
01/08/24	CHECKCARD 0107 BB* UNDERDOGS CANTINA HTTPSBBOT.MENCA 24011344007000007761390	-74.56
01/08/24	MOBILE PURCHASE 0107 LAZ PARKING 680208-DP SAN FRANCISCOCA	-10.00
01/09/24	MOBILE PURCHASE 0107 DENNY'S #9622 LIVERMORE CA	-60.58
01/09/24	CHECKCARD 0108 APPLE.COM/BILL 866-712-7753 CA 24692164008108842503856 RECURRING	-9.99
01/09/24	MOBILE PURCHASE 0108 CHEVRON 0212470 TRACY CA	-30.52
01/10/24	CHECKCARD 0107 THE CLANCY, AUTOGRAPH SAN FRANCISCOCA 24692164009 1095641 93065	-358.61
01/10/24	CHECKCARD 0110 CHEVRON 0212470 TRACY CA 24692164010109922776970	-46.02
01/12/24	Z VAPE N DISCO 01/12 #000002927 MOBILE PURCHASE 245 EAST 11TH STR TRACY CA	-25.00
01/12/24	FOOD MAXX #490 01/12 #000845946 PURCHASE FOOD MAXX #490 TRACY CA	-7.87
01/16/24	PMNT SENT 0113 CASH APP*DONALD WILSON* 800-9691940 CA 24793384013001491975084	-20.00
01/16/24	TRACY MARKET 01/13 #000348400 MOBILE PURCHASE 15 E GRANT LINE R TRACY CA	-26.07

The studio edit loop

Local Region Reruns: `run_tsr_on_bbox` and `run_ocr_on_bbox` expose surgical fixes without reprocessing the entire page.

Auto-Persist Editing: Complete grid mutation tools (split, merge, find/replace) execute rapidly with a 350ms debounced auto-save.

Dynamic Crop Padding: Base 4% padding with a cross-axis cap of 20% prevents runaway model expansion on thin or wide bounding boxes.

Agent P-DF 00063.jpg

PAGE 1

Consolidated Statement of Comprehensive Income
for the year ended 31 December 2020

Table 1

	2020			2019		
	Normalised Significant results	IFRS	IFRS	Normalised Significant results	IFRS	IFRS
Profit for the year	3,717	(2,921)	796	9,068	(2,968)	6,100
Items that may be reclassified subsequently to profit or loss						
Net fair value of cash flow hedges	(556)		(536)	(525)		(525)
Translation difference	(378)		(378)	236		236
Total comprehensive income (loss) for the year	2,783	(2,921)	(138)	8,779	(2,968)	5,811
Attributable to						
Owners of the Company	1,950	(2,921)	(971)	7,664	(2,968)	4,696
Non-controlling interests	833		833	915		915
	2,783	(2,921)	(138)	8,779	(2,968)	5,811

Confidence sensitivity: 0%

Selection padding: 0px

Grid Text Heatmap

Pointer Marquee Draw Table

Run TSR + OCR Delete

Add Cell Delete Cell

Pages: 1/1 Speed: 3.4 sec/page Errors: 0

Back / New Upload Expert CSV Export Excel Export HTML

Viewing: Page 1 Table: Table 1

	Note	2020			2019	
		Normalised Significant results		Normalised Significant items*		
		IFRS		IFRS		
Profit for the year	3,717	(2,921)	796	9,068	(2,968)	6,100
Items that may be reclassified subsequently to profit or loss						
Net fair value of cash flow hedges	(556)		(536)	(525)		(525)
Translation difference	(378)		(378)	236		236
Total comprehensive income (loss) for the year	2,783	(2,921)	(138)	8,779	(2,968)	5,811
Attributable to						
Owners of the Company	1,950	(2,921)	(971)	7,664	(2,968)	4,696
Non-controlling interests	833		833	915		915
	2,783	(2,921)	(138)	8,779	(2,968)	5,811

Document Tools and Granular Agency

AI imperfections require surgical tools. The studio provides dynamic confidence filtering to hide hallucinated tables, adjustable padding sliders to control extraction margins, and multiple interaction modes to capture data exactly where the pipeline failed.

Confidence sensitivity

0%

Selection padding

0px

Grid

Text

Heatmap

Pointer

Marquee

Draw Table

Run TSR + OCR

Delete

Add Cell

Delete Cell

Pointer	Select, drag, and resize bounding boxes manually.
Marquee	Drag a region to instantly run localized OCR copy.
Draw Table	Manually trigger full structural recognition (TSR) and OCR on any drawn region.

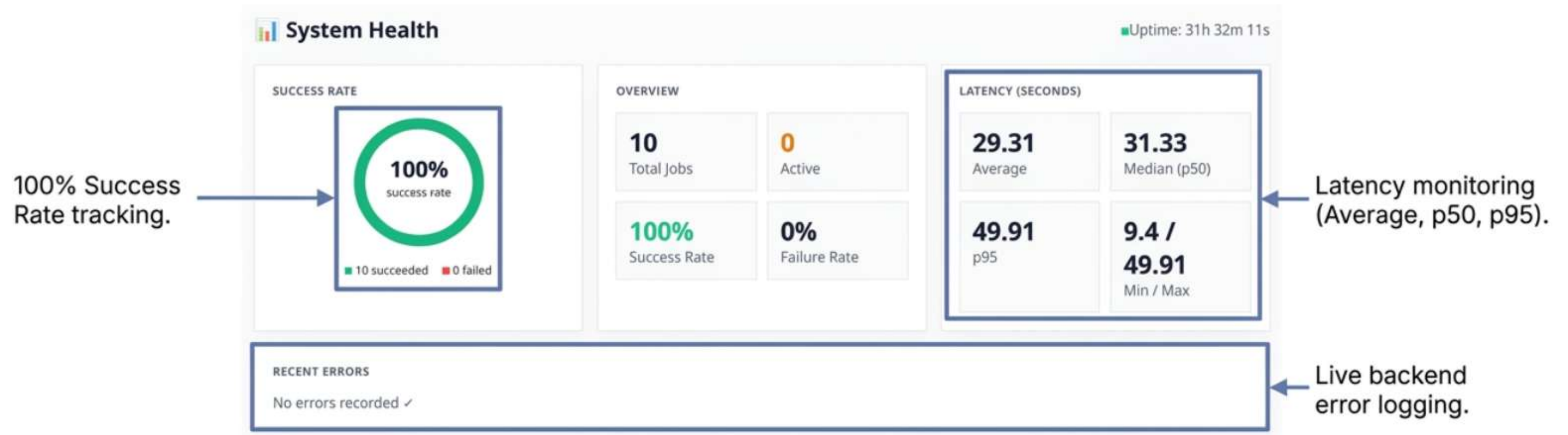
Deterministic semantics and typographic control.

Heuristics View: Our deterministic Python engine classifies table roles (e.g., StubHeader, SuperRowHeader) mathematically. The operator can instantly visualize these classifications via role coloring to verify semantic accuracy.

Grid Editor Capabilities:

- Selection-aware Find and Replace across complex tables.
- Full-row / full-column merge behavior.
- Typography cycling and custom symbol palettes for highly specialized financial or medical extraction.

Fleet management and system health



Built for single-process, local-first enterprise deployment.
Complete operational visibility into uptime, throughput, and system latency.

Export Semantics and Business Handoff

The ultimate goal of the system is structured data. The export layer translates the pipeline's spatial geometry and the user's manual grid corrections into native business formats, preserving complex layouts like spanned cells directly into native Excel formatting.

Format	Behavior
JSON	Full annotation document payload.
HTML	One distinct table snippet per detected grid.
CSV	All tables mapped to a single contiguous stream with separator rows.
XLSX	Multi-sheet generation with spatial spanning cells perfectly preserved as Excel merged cells.

The Architecture of Reliability

Agent P-DF accepts that machine learning will never achieve absolute perfection on unpredictable, real-world documents. By architecting a robust AI foundation and bridging it to a flawless editing environment, it delivers the only metric that matters: 100% accurate data extraction.

Discriminative AI

62.5M parameter pipeline, trained on hard negatives for real-world robustness.

Deterministic Heuristics

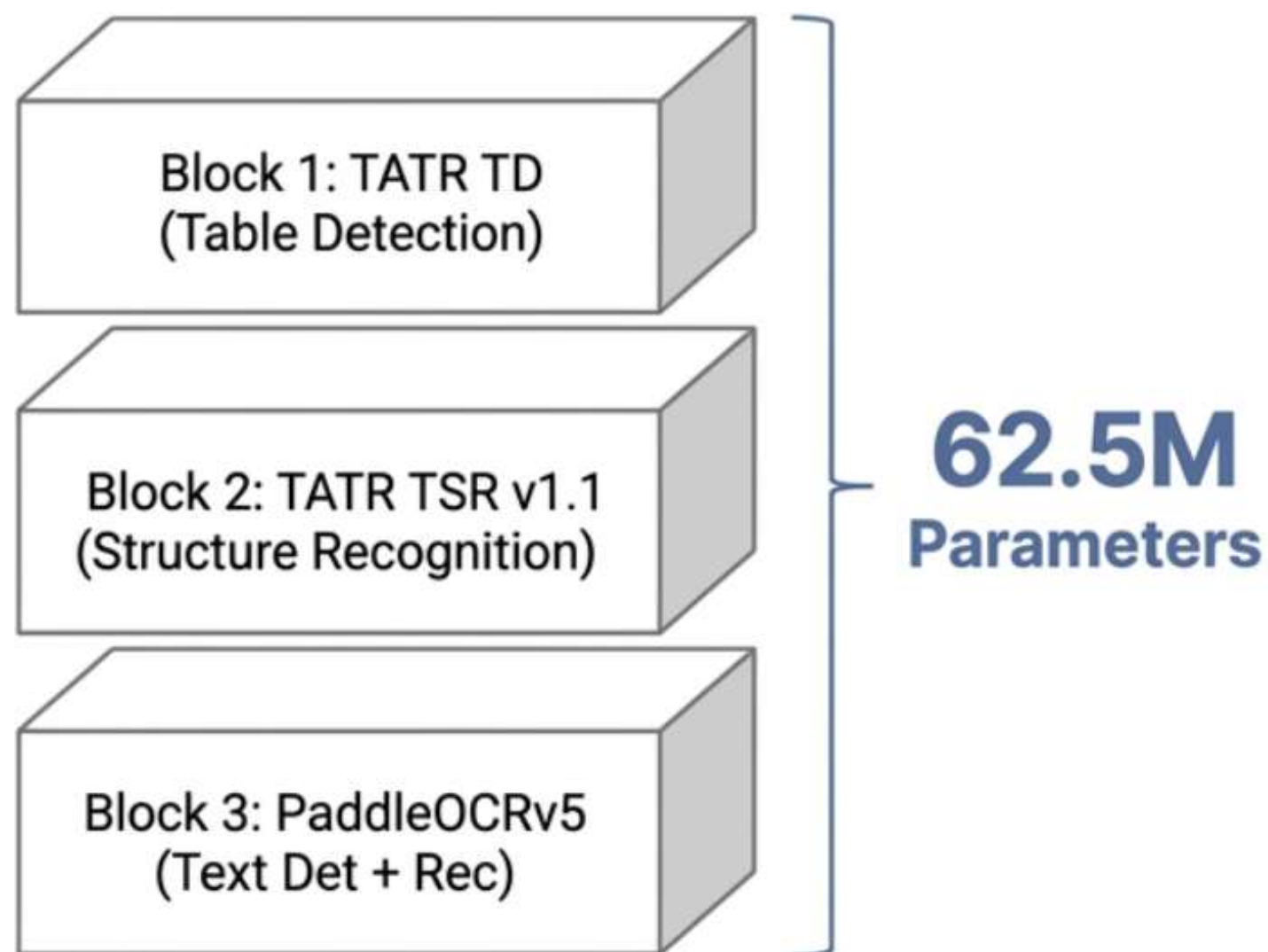
Algorithmic bypasses and mathematical boundaries that sanitize raw predictions.

Flawless UX

A split-screen studio that grants total agency to the user for final verification.

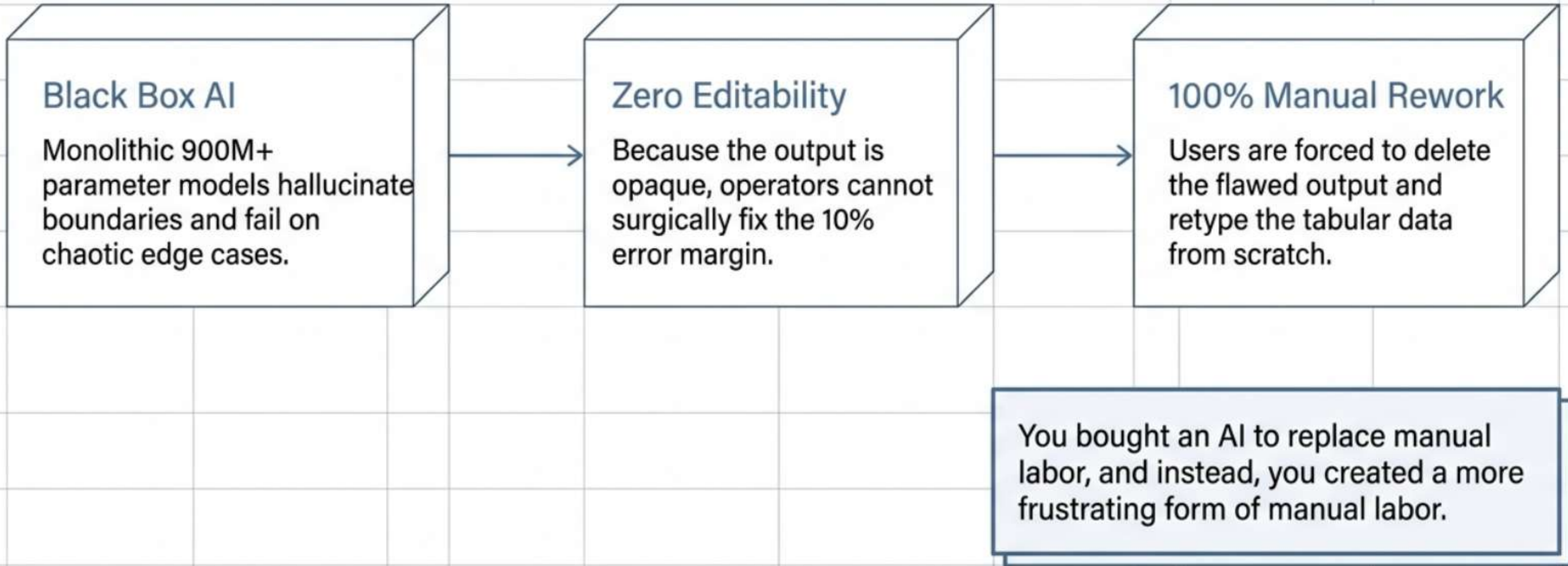
Product *As* Business

Engineered for the lightweight gap.

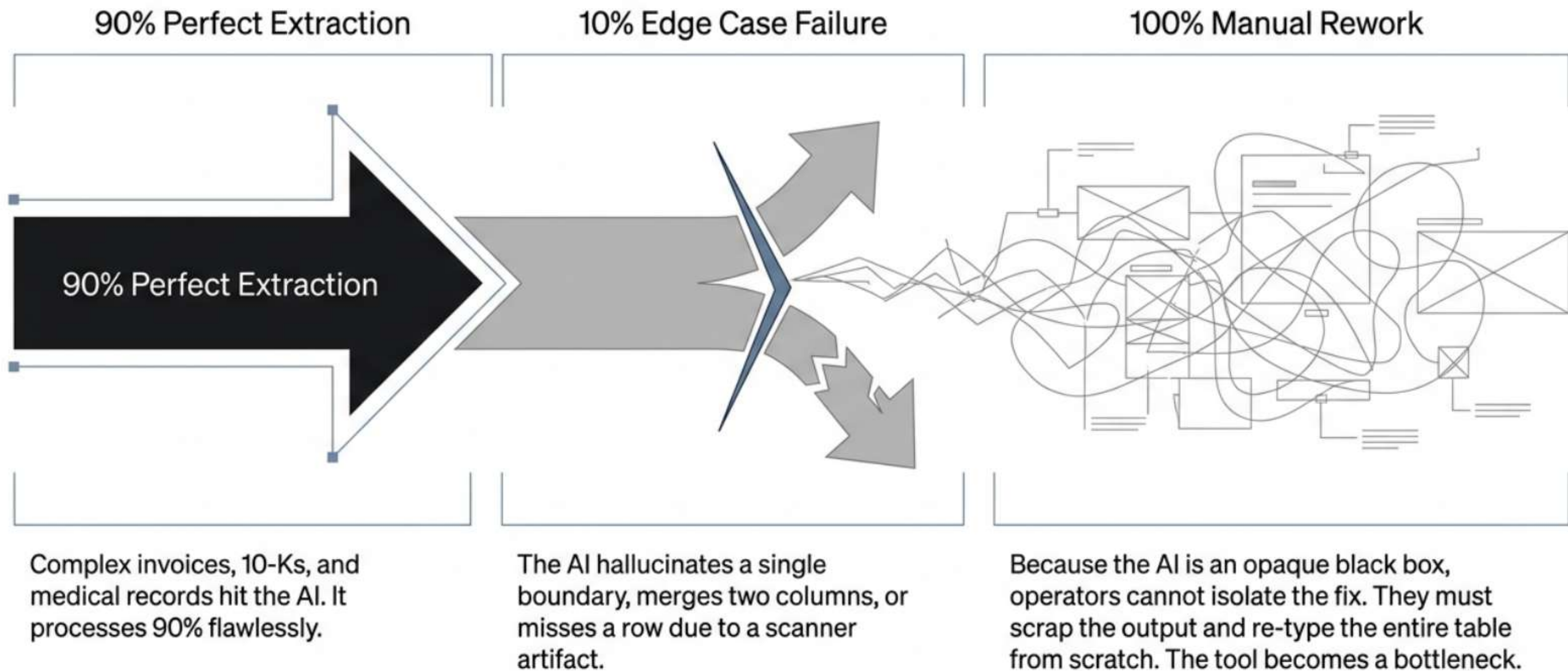


- Runs flawlessly on local CPUs.
- Zero reliance on massive GPU farms.
- Fractional compute costs compared to monolithic VLMs or cloud APIs.

Zero-touch automation is an expensive lie.



The 10 percent error rate trap.



The studio is the product

The Source: AI-driven geometric mapping.

BANK OF AMERICA

DONALD WILSON II | Account # 2021 7740 2344 | November 14, 2023 to January 16, 2024

Withdrawals and other subtractions - continued

Date	Description	Amount
01/06/24	MOBILE PURCHASE OTOB Reef Parking Miami FL	-40.00
01/06/24	CHECKCARD 0100 B8P UNDERROOS CANTINA HTTPSBBOT MENCX 24011344007000007761390	-18.00
01/06/24	MOBILE PURCHASE 0107 LAZ PARKING 688256 OP SAN FRANCISCOCA	-10.50
01/06/24	MOBILE PURCHASE 0107 DENNY'S #9522 LIVERMORE CA	-46.30
01/06/24	CHECKCARD 0176 APPLE CONEILL 886-772-7755 CA 24692164008108842503556 RECURRING	-4.40
01/06/24	MOBILE PURCHASE 0107 CHEVRON C212470 TRACY CA	-33.50
01/06/24	CHECKCARD 0117 THE CLANCY, AUTOGRAPH SAN FRANCISCOCA 246921640091095641 93065	-38.51
01/06/24	CHECKCARD 0110 CHEVRON 0212470 TRACY CA 24692164010169922776970	-43.50
01/06/24	Z VAPE N DISCO 0113 MOBILE PURCHASE 355 EAST 11TH STR TRACY CA	-17.30
01/06/24	FOOD MAXX #490 01/12 #000845945 PURCHASE FOOD MAXX #490 TRACY CA	-15.00
01/06/24	PMNT SENT 0113 CASH APP*DONALD WILSON* 800-9691940 CA 24793384013001191975064	-35.10
01/06/24	TRACY MARKET 01/13 #000338400 MOBILE PURCHASE 15 E GRANT LINE E TRACY CA	-35.40
01/06/24	MOBILE PURCHASE 0115 MODOVALLOS 73522V TRACY CA	-32.16
01/06/24	PMNT SCAT 0111 CASH APP*DONALD WILSON* 800-9691940 CA 24793384013001191975064	-33.50
01/06/24	PURCHASE 0111 ACZ HARDWARE CORPORATION 680-REI-0900A	-12.30
01/06/24	CHEVRON CARD 0113 CHEVRON 012470 TRACY CA 24692164010169922776970	-44.50

Total ATM and debit card withdrawals: \$3,100.00

If there were tax charges paid or debited to your account to cover the electronic transactions shown at 01/16/23, the amount(s) indicated here have been paid.

Page 2

Donal Wilson

Date	Description	Amount
01/16/24	Zelle payment to MELIA Certif-nations	365.00

The Environment:
Deterministic
structured grid
editor.

Agent P-DF is not a parser; it is a human-in-the-loop workspace.

The environment provides split/merge tools, confidence heatmaps, and zero-latency corrections in one continuous loop.

Built for the realities of financial
and scientific data.

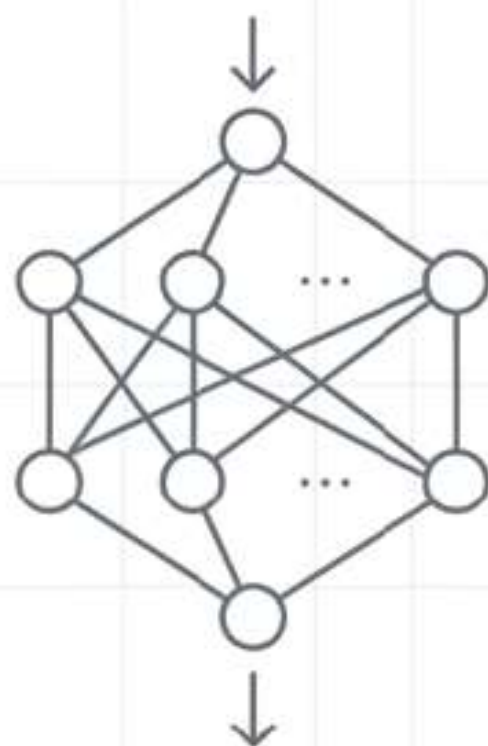
Standard OCR constantly misinterprets the highly specific symbols found found in enterprise documents.

The image shows a 'Quick Symbols' dialog box overlaid on a spreadsheet. The dialog box has a title bar with a close button (X) and a subtitle 'Click to copy to clipboard'. Below the subtitle is a text input field labeled 'Add custom text/symbol...' and an 'Add' button. The main content area is divided into sections: 'CUSTOM SYMBOLS' (with a scrollable list containing 'EGP'), 'FINANCIAL' (with buttons for currency symbols: \$, €, £, ¥, ₹, P, W, ₮, ¢, %, ‰), 'SCIENTIFIC' (with buttons for Greek letters: θ, π, σ, μ, Δ, Ω, α, β, γ, λ, ∞), and 'MATH' (with buttons for mathematical symbols: ∞, ≈, ≤, ≥, √, ∫, Σ, °, ±). The background spreadsheet has columns for 'Date', 'Description', and 'Amount'. The 'Description' column contains various transaction details, and the 'Amount' column contains numerical values, some with negative signs.

- Integrated typography panel mapped directly to local storage.
- One-click insertion for Financial (\$, €, £, ¥), Scientific (θ , π , μ), and Math symbols.
- Instant find/replace across the entire structured table matrix.

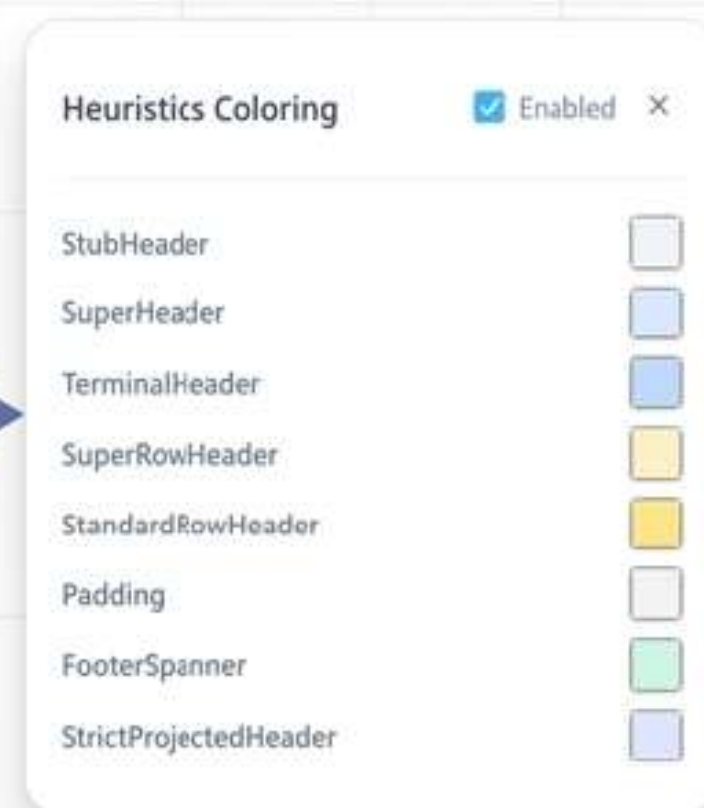
Splitting the brain from the math

The Neural Network



Handles pure geometry.
Finds lines and bounding boxes.
Kept intentionally small (62.5M
parameters) to avoid hallucination.

The Deterministic Mastermind



Python heuristics handle semantics,
spanning cells, and text anchoring.
Code does not guess; it calculates.

Enterprise compliance by design. Total data privacy and legal immunity.

100% Local On-Premise

At just 62.5M total parameters, the entire pipeline runs locally. No cloud API calls. No sensitive data ever leaves the building.

GPL-Free Architecture

Deliberately abandoned the YOLO architecture to strip out toxic open-source licenses, ensuring **zero IP risk** or legal exposure for enterprise deployment.

CPU-Capable Compute

PIPELINE_DEVICE=cpu supported out of the box. Drastically lowers operational overhead—no massive GPU server racks required.

Operational oversight in real-time.

Designed for production environments, the system features an in-memory health dashboard that polls critical metrics without complex external dependencies.



Live Uptime &
Total Job Counts

Success vs.
Failure Rates

p50 and p95
Latency Statistics

Stop trusting black-box AI with enterprise data.

Viewing: Page 1 Table **Table 1** ▾

↶ ↷

🔍

✂

📌

📏

🗑

🔄

H

T

🎨

🔗

Σ

✂

📄

Find...

Replace...

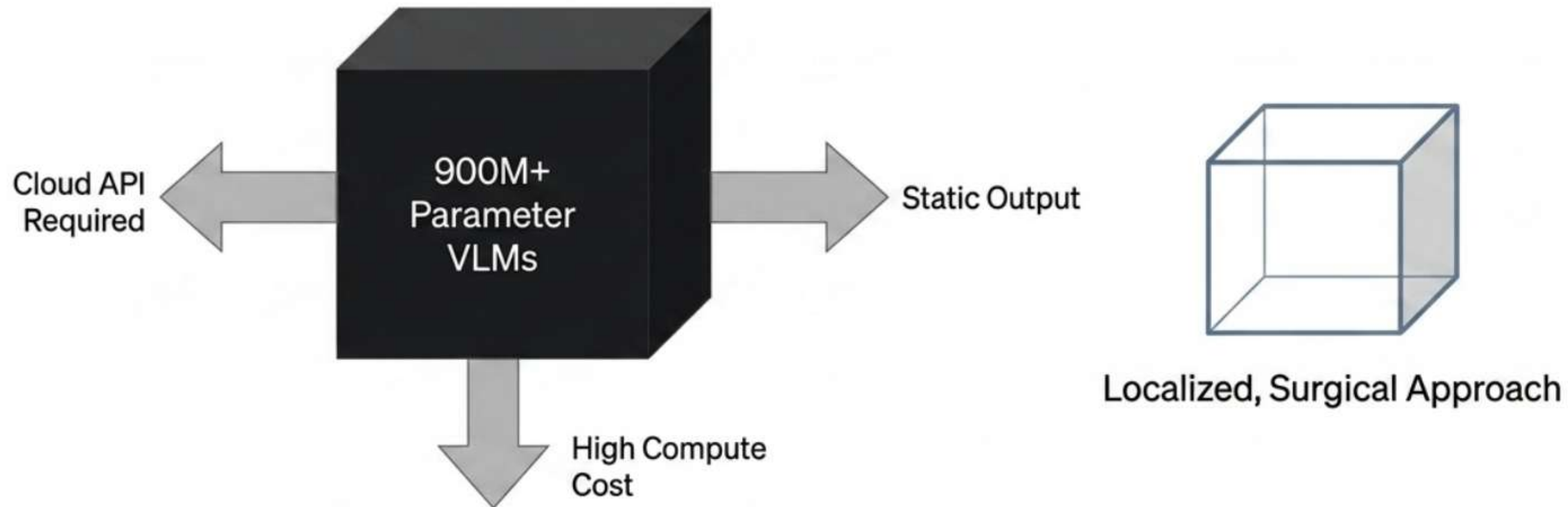
Replace All

	ATM and debit card subtractions - continued	
Date	Description	Amount
01/08/24	MOBILE PURCHASE 0106 Reef Parking Miami FL	-400.
01/08/24	CHECKCARD 0107 BB* UNDERDOGS CANTINA HTTPSBBOT.MENCA 24011344007000007761390	-725.

Every other tool gives you a noisy output and wishes you luck. Agent P-DF gives you a studio where you can fix it in seconds—**locally, privately, and surgically.**

Agent P-DF. Built for the final mile.

The fallacy of monolithic vision models.



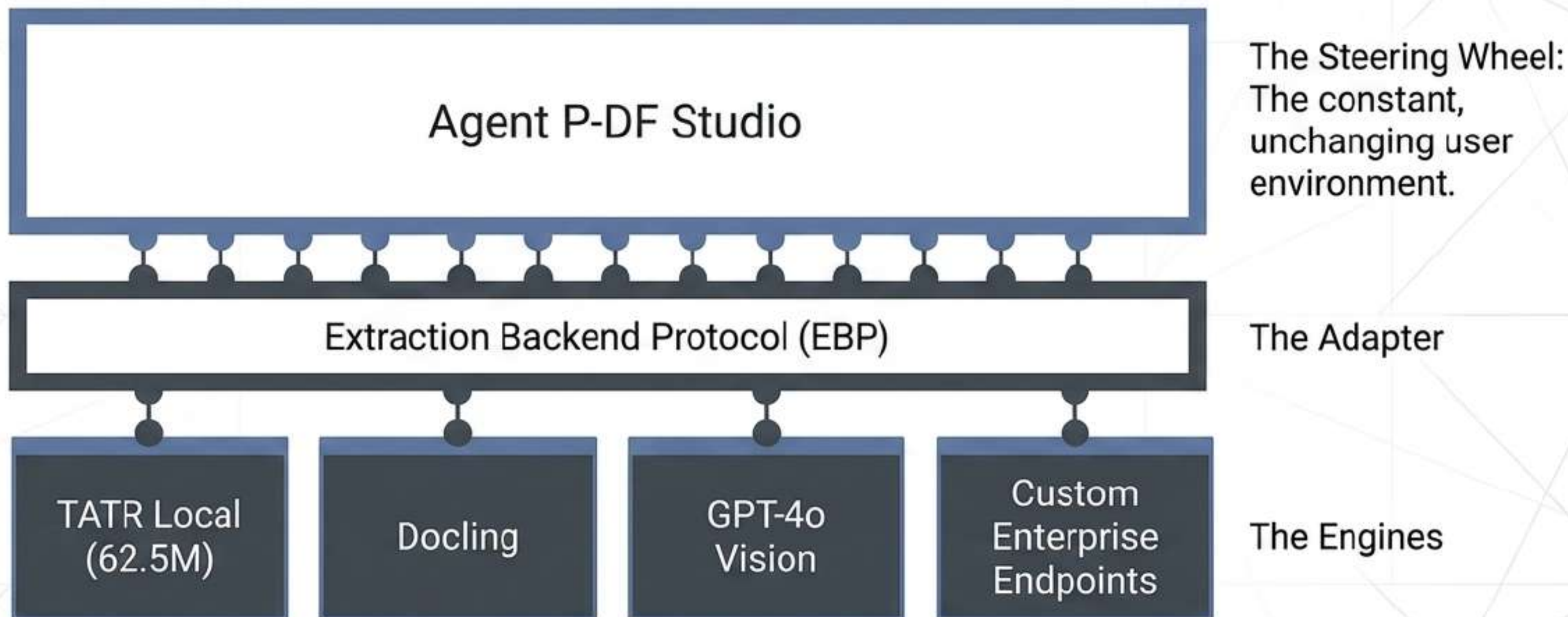
- Privacy leaks: Sending sensitive financial or healthcare data to third-party cloud APIs.
- Over-engineered bloat: Requiring massive GPU farms for simple document processing.
- The Yes-Man problem: Trained on sanitized data, causing them to confidently hallucinate tables over flowcharts.

Bypassing the market extremes.

	Deployment	Parameter Footprint	Output Editability	Privacy / Compliance
Cloud APIs like AWS Textract	Public Cloud	N/A	Zero (Static API output)	High Risk (Data leaves premise)
Monolithic VLMs like PaddleOCR VL 1.5	Cloud or Heavy GPU	900M+	Zero (Markdown/Text)	Variable
Agent P-DF	✓ 100% Local	✓ 62.5M	✓ Surgical Studio UI	✓ Total (Air-gapped capable)

AI handles geometry, deterministic code handles semantics, and UI handles the residual error.

Bring your own intelligence



- Models are commodities. We own the workflow.
- Plug in any intelligence; the human editing experience never changes.

The bottom line.



Agent POF

**Every other tool gives you a noisy
output and wishes you luck.
We give you a surgical studio
where you fix it in seconds.**

Scan Me

