

SHOPIFY E-COMMERCE OPERATIONS

CATALOG QUALITY AUTOMATION

A code-first, AI-assisted workflow with human-review gates for turning inconsistent supplier inputs into pre-import validated Shopify CSV exports.

1,775

STORE-PRODUCT LISTINGS

Distinct by catalog and Handle

24,395

VARIANT ROWS

Rows with unique Variant SKU

15,227

PRODUCT-IMAGE RECORDS

Rows with Image Src

13

FINAL GENERATED BATCHES

Across two anonymized catalogs

CODE-FIRST

AI-ASSISTED

HUMAN-REVIEW GATES

Thai Anh Tu

Catalog scale turns small errors into operational risk

Supplier storefronts vary in structure, language, currency, option logic, and gallery behavior. Manual copying multiplies defects across every Color x Size combination.

01

Inconsistent sources

Different platforms and page structures produce uneven titles, prices, options, taxonomy, and images.

02

Variant explosion

One product can expand into many SKU-bearing rows across color and size combinations.

03

Image mismatch

A wrong color-image relationship can make an otherwise valid product commercially unusable.

04

Silent defects

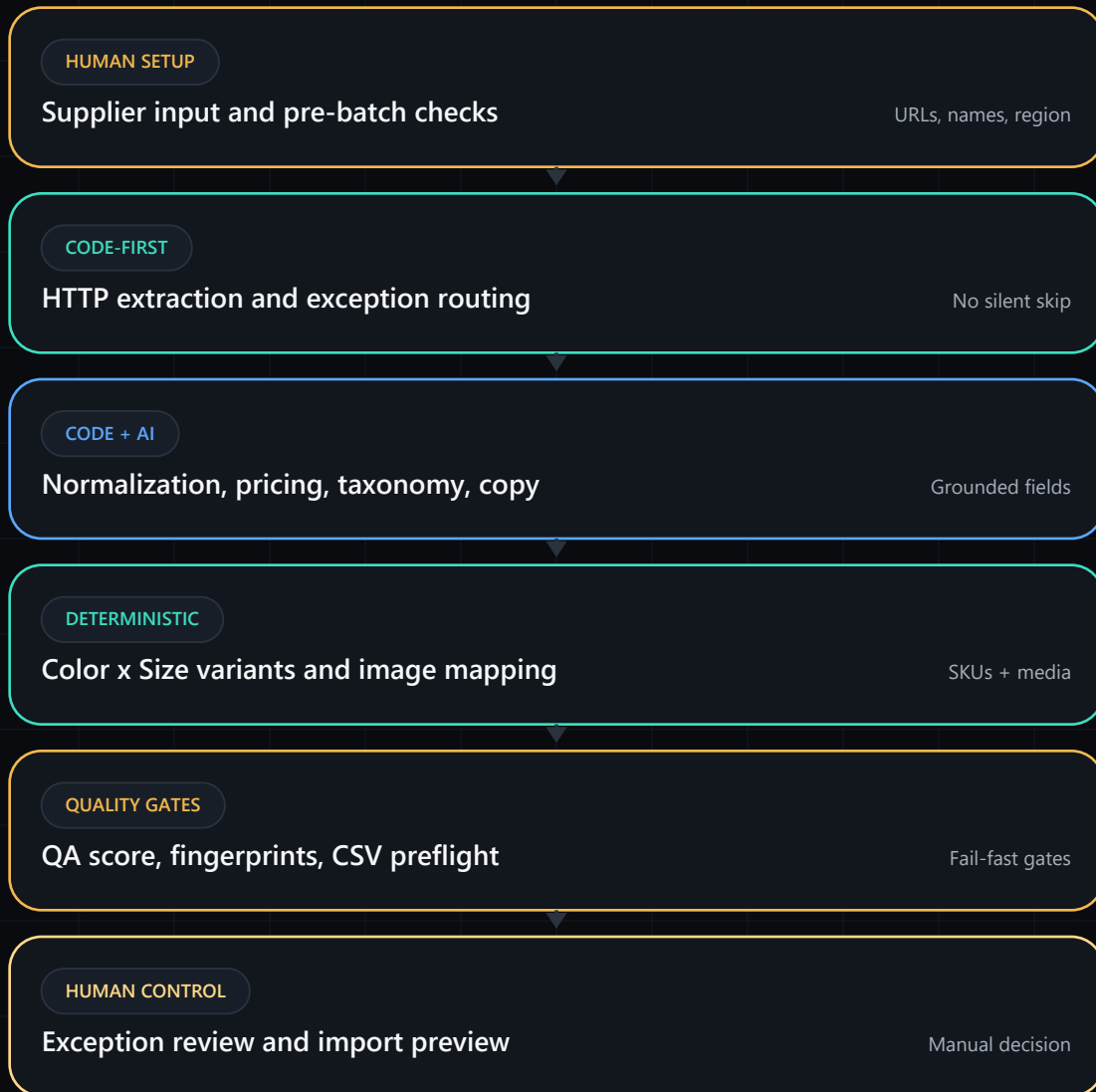
Duplicates, stale files, leaked supplier terms, and malformed data can survive spreadsheet-only repair.

MY OPERATING RESPONSE

I designed and operated a configurable workflow that converts heterogeneous supplier inputs into structured Shopify catalog data, surfaces exceptions, and blocks release when critical checks fail.

From supplier input to import-ready Shopify CSV

Repeatable transformation is automated. Ambiguous extraction, warnings, exceptions, final upload, and post-import checks stay visible to a human.



OUTPUT: PRE-IMPORT VALIDATED SHOPIFY CSV + EXCEPTION REPORTS

One input row becomes a complete product model

	A	B	C
1	country	product_name	link
2	GB	Harbor Linen Shirt	https://supplier-a.example.invalid/products/harbor-linen-shirt
3	US	Aster Wrap Dress	https://supplier-b.example.invalid/products/aster-wrap-dress
4	AU	Cove Knit Cardigan	https://supplier-c.example.invalid/products/cove-knit-cardigan

SYNTHETIC INPUT: COUNTRY | PRODUCT NAME | AUTHORIZED SOURCE LINK



IDENTITY

Handle, title, vendor, taxonomy

OPTIONS

Color and Size combinations

COMMERCE

SKU, price, compare-at, status

SYNTHETIC OUTPUT EXCERPT

Handle	Color	Size	Variant SKU	Price	Image
harbor-linen-shirt	Navy	S	DEMO-HLS-001	49.95	Mapped
harbor-linen-shirt	Navy	M	DEMO-HLS-002	49.95	Mapped
harbor-linen-shirt	Ivory	S	DEMO-HLS-003	49.95	Mapped
harbor-linen-shirt	Ivory	M	DEMO-HLS-004	49.95	Mapped
aster-wrap-dress	8		DEMO-AWD-001	69.95	Fallback
aster-wrap-dress	10		DEMO-AWD-002	69.95	Fallback

Synthetic demo illustrating schema and row relationships. Products, SKUs, prices, markets, and URLs are fabricated. No client data is shown, and this sample was not imported.

Generated output at catalog scale

An evidence audit selected the latest non-overlapping final CSV cohorts and removed byte-identical backups, superseded versions, and corrective subsets before counting.

1,775

STORE-PRODUCT LISTINGS

Distinct by catalog and Handle

24,395

SKU-BEARING ROWS

All SKUs unique in selected set

15,227

PRODUCT-IMAGE RECORDS

Rows containing Image Src

37,847

PHYSICAL DATA ROWS

One consistent 41-column schema

ROW RECONCILIATION

$$24,395 + 15,227 - 1,775 = 37,847$$

Variants + Image Src - shared first rows

13 GENERATED BATCHES

2 ANONYMIZED CATALOGS

41 COLUMNS

Scope note: these figures describe locally generated files. They do not claim client receipt or successful Shopify import.

Designed to fail visibly - not silently

The release path exposes exceptions and blocks critical defects instead of relying on a final spreadsheet glance.

RELEASE GATES

1

Pre-batch validation

Required fields, duplicate links, workflow readiness

2

Raw-data QA

Content, price, size, image, taxonomy, encoding

3

Weighted product score

Product and batch thresholds

4

Fingerprint and CSV hash

Blocks stale, changed, or forced output

5

Pre-import and human review

P0 blockers, warnings, exceptions, preview

72

RAW-QA RULE IDS

52 ERROR + 20 WARN IDs present

244

TEST FUNCTIONS

Across 28 source modules

0

DUPLICATE SKUS

Within 24,395 selected variants

STRUCTURAL AUDIT OF THE SELECTED EXPORT SET

- One consistent 41-column schema
- No malformed rows or blank handles
- 15,227 unique catalog-Handle-Image Src relationships

Source counts were statically audited. The current 244-test suite was not freshly executed, and these controls do not imply zero defects.

Operations, architecture, and quality control

I combined catalog operations judgment with AI-assisted implementation to build a repeatable workflow, while keeping human control over ambiguous data, exceptions, and release decisions.

CONTRIBUTION AREAS

- Store-scoped configuration and batch workflow
- Supplier extraction and normalized product schemas
- Shopify taxonomy, pricing, and variant generation
- Image handling and color-to-image resolution
- QA rules, scoring, fingerprints, and preflight
- Exception queues, reports, and operating documentation

AUTOMATION BOUNDARY

DETERMINISTIC

Normalization, pricing, variants, CSV, QA, hashing

AI-ASSISTED

Difficult-page residue and grounded product-copy drafting

HUMAN CONTROL

Configuration, exception decisions, import, post-import checks

RECRUITER KEYWORDS

SHOPIFY

E-COMMERCE

CATALOG MANAGEMENT

PRODUCT DATA

WORKFLOW AUTOMATION

QUALITY ASSURANCE

Role framing: workflow design, catalog operations, AI-assisted implementation, QA architecture, and human review.



THAI ANH TU

CATALOGS. CONTROLLED. READY TO SCALE.

Shopify E-commerce Operations | Catalog and Product Data | Workflow Automation

OPERATIONAL VALUE

REPEATABILITY

A defined path from source input to structured export.

TRACEABILITY

Fingerprints, hashes, reports, and visible exceptions.

DEFECT PREVENTION

Layered controls catch known catalog risks earlier.

HUMAN CONTROL

Warnings, edge cases, upload, and final checks stay human-owned.

OPEN TO REMOTE FULL-TIME OR CONTRACT OPPORTUNITIES

Shopify operations, catalog and product-data management, e-commerce workflow automation, and creative operations.

Anonymized case study. Metrics are scoped to selected locally generated exports across two catalogs.
No client identities, production data, credentials, revenue claims, or unverified speed claims are included.