



DOCUMENT & DATA PIPELINE

Document Processing and Metrics Pipeline
Representative work sample | Sanitized composite

OVERVIEW

WS03
5-7 weeks
Forge Honor Assurance



Document Processing and Metrics Pipeline

Representative sample engagement for replacing spreadsheet-heavy document handling.

ENGAGEMENT FIT

Teams using manual copy/paste, PDF review, shared drives, email handoffs, and ad hoc spreadsheets.

SCENARIO

A business unit receives structured forms, semi-structured PDFs, email notes, and spreadsheet updates. The team manually extracts fields, checks for missing information, flags exceptions, and compiles weekly metrics. The process is slow, hard to audit, and produces inconsistent reporting.

CONSTRAINTS

- Source documents must remain traceable to each extracted field.
- The pipeline must support manual correction when extraction is uncertain.
- The first release needs to create useful operational reporting without replacing the buyer's system of record.
- The workflow must tolerate imperfect documents rather than failing silently.

This is a representative composite report showing the type of work product Forge Honor Assurance can provide. It is not a client case study.

REPORT DATA

SAMPLE ID
WS03

CATEGORY
Document & Data Pipeline

TIMEFRAME
5-7 weeks

BEST FOR
Teams using manual copy/paste, PDF review, shared drives, email handoffs, and ad hoc spreadsheets.

USE NOTE

Representative composite sample. No client PII, proprietary environment detail, or confidential facts are included.

Final scopes, controls, deliverables, and timelines are confirmed in writing per engagement.



DOCUMENT & DATA PIPELINE

Document Processing and Metrics Pipeline
Representative work sample | Sanitized composite

DELIVERY PLAN

WS03
5-7 weeks
Forge Honor Assurance



Scope and Delivery Plan

A practical work plan with written scope, explicit boundaries, and deliverables the buyer can inspect.

SCOPE

- Define source types, required fields, validation rules, exception categories, and reporting needs.
- Build extraction and normalization workflow for approved document classes.
- Create validation and exception handling so humans review uncertain or incomplete records.
- Generate dashboard-ready output tables and weekly quality metrics.
- Deliver deployment notes, operator runbook, and known limitations.

TIMELINE

PHASE	FOCUS	EXPECTED WORK
Week 0	Discovery	Inventory source documents, field definitions, data owners, and reporting outputs.
Week 1	Schema and rules	Define normalized schema, validation rules, exception categories, and confidence thresholds.
Weeks 2-4	Pipeline build	Implement ingestion, extraction, normalization, validation, and export paths.
Week 5	Dashboard and QA	Produce metrics views, exception review flow, and quality checks.
Weeks 6-7	Pilot and handoff	Pilot with real samples, tune rules, document operating cadence, and train reviewers.

DELIVERABLES

- **Data dictionary:** Field definitions, source mapping, validation rules, owners, and example values.
- **Processing pipeline:** Repeatable ingestion, extraction, normalization, validation, and export workflow.
- **Exception queue:** Records requiring review, reason codes, confidence notes, and reviewer actions.
- **Reporting tables:** Dashboard-ready outputs for volume, quality, turnaround, and unresolved exceptions.
- **Runbook:** Operator steps, failure handling, correction workflow, and maintenance notes.



DOCUMENT & DATA PIPELINE

Document Processing and Metrics Pipeline
Representative work sample | Sanitized composite

SAMPLE OUTPUT

WS03
5-7 weeks
Forge Honor Assurance



Sample normalized output

The pipeline should produce records that can be traced, corrected, and reported.

FIELD	EXAMPLE	CONTROL
source_id	DOC-2026-0148	Links back to original file and processing run.
request_type	Eligibility update	Controlled taxonomy, not free-form text.
due_date	2026-06-14	Date parser plus validation against business rules.
missing_fields	approval_signature	Drives exception queue and reviewer assignment.
review_status	Needs review	No downstream export until required fields pass.

HANDOFF

- The buyer receives a pipeline that can be operated, monitored, and corrected without relying on hidden manual steps.
- The runbook includes what to do when source formats change, fields fail validation, or output exports break.
- Future phases can add deeper integrations, more document classes, or stricter workflow automation.



DOCUMENT & DATA PIPELINE

Document Processing and Metrics Pipeline
Representative work sample | Sanitized composite

SAMPLE OUTPUT

WS03
5-7 weeks
Forge Honor Assurance



Sample weekly metrics

Operational reporting should show whether the workflow is getting healthier.

METRIC	EXAMPLE	WHY IT MATTERS
Documents processed	412	Tracks volume and capacity.
Auto-pass rate	68%	Shows how many records avoid manual correction.
Exception rate	21%	Highlights document or process quality issues.
Median review age	9.4 hours	Shows backlog pressure.
Top exception	Missing approval signature	Identifies process fix opportunity.

BOUNDARIES

- The first release should not replace the system of record unless explicitly scoped.
- Extraction quality depends on source document consistency and approved review thresholds.
- The engagement does not decide records policy, retention law, or regulatory interpretation.