

URL LEDGER

# Audit Intake & Data Access Workbook

How to prepare a client, gather the right inputs, assign stakeholders, and convert a 45-day portfolio repricing audit into an always-on URL asset ledger.

**CORE PROMISE**

We do not need a perfect warehouse on day one. We need enough structured truth to identify value leakage, reprice the URL portfolio, and define what humans or agents are allowed to change.

Prepared for pilot partner onboarding, audit delivery, and implementation handoff

# 1. Purpose of This Workbook

This workbook converts interest into audit readiness. It gives sales, delivery, technical, and client-side stakeholders a shared intake path before the 45-day URL Portfolio Repricing Audit begins.

- Clarify the business problem: decay, waste, conversion leakage, attribution uncertainty, or agent-readiness risk.
- Define the unit of account: which domains, subdomains, folders, templates, and clusters belong in scope.
- Collect the minimum viable data needed to score URL assets across the 13 structural variables.
- Protect security and confidence through read-only access, exports, evidence logging, and clear retention rules.
- Set a clean path from audit findings to ledger installation, action backlog, governance, and recurring operations.

## OPERATING PRINCIPLE

The audit should feel like capital allocation, not content cleanup. Every input should help answer: what exists, what is it worth, what is decaying, what should be protected, and what should happen next?

## 2. Intake Summary: What We Need Before Day 1

Need	Input	Why it matters
Required	Domain / property scope	Primary domain, subdomains, regions, folder boundaries, exclusions, and any recent migrations.
Required	Sitemap or crawl permission	XML sitemap, CMS export, or permission to crawl the public site.
Required	Google Search Console	Page/query impressions, clicks, CTR, position, coverage, and indexation signals.
Required	Analytics	GA4 or equivalent landing page data, sessions, events, conversions, and traffic source context.
Required	Conversion definition	One to three primary conversion events or revenue outcomes used to define yield.
Recommended	Revenue truth source	CRM closed-won, pipeline, ecommerce orders, payments, lead value proxy, or RPM model.
Recommended	CMS / content export	Publish date, update date, author, template, content type, owner, and status.
Optional	Paid / lifecycle data	Google Ads, Meta, LinkedIn, Klaviyo, HubSpot, Marketo, or campaign data for assisted value.
Optional	Warehouse / BI	BigQuery, Snowflake, Redshift, Looker, Power BI, or internal reporting extracts.

### 3. Stakeholder Map

Role	Responsibility	Typical person
Executive sponsor	Approves business objective, scope, and readout audience.	CMO, CFO-minded CMO, VP Growth, Head of Digital
Data owner	Grants analytics, search, CRM, and warehouse access or exports.	Analytics lead, RevOps, BI owner
Technical owner	Answers crawl, canonicals, redirects, templates, indexation, and CMS questions.	SEO lead, web lead, engineering
Content owner	Explains content strategy, ownership, publish history, and protected pages.	Content lead, editorial, product marketing
Revenue owner	Defines conversion value, pipeline stages, revenue attribution, and acceptable proxies.	RevOps, sales ops, finance
Governance approver	Defines what cannot be touched, what requires approval, and what agents may only recommend.	Legal, brand, compliance, product owner

#### CLIENT BURDEN TARGET

The ideal intake requires 60-90 minutes of stakeholder time across the full audit cycle, excluding access setup. Exports are acceptable if the client does not want live read-only connections.

### 4. Scope Definition Worksheet

Scope item	Client notes
Primary domain(s)	
Subdomains included	
Subdomains excluded	
Countries / regions / languages	
Folders or sections in priority scope	
Known recent migrations, redesigns, CMS changes, or template changes	
Known high-value page groups	
Known problem areas or suspected decay clusters	
Pages requiring no-touch / high-approval treatment	

## 5. Business Objective and Monetization Mapping

A URL asset can create value through revenue, pipeline, assisted conversion, demand generation, retention, enablement, risk reduction, or brand authority. The intake should define which value model applies before scoring begins.

Business model	Value definition	Common truth source
Ecommerce	Orders, gross revenue, margin, AOV, repeat purchase, product category value.	Shopify, WooCommerce, Stripe, GA4 purchase events
Lead generation	Qualified leads, demos, form submissions, calls, opportunity creation, closed-won pipeline.	HubSpot, Salesforce, Zoho, Dynamics, call tracking
SaaS	Trial starts, demos, signups, activation, expansion, assisted pipeline, product-led conversions.	CRM + product analytics + GA4
Publisher / media	RPM, ad revenue, affiliate clicks, subscriptions, newsletter signups.	Ad server, analytics, affiliate platform
Local / services	Calls, bookings, quote requests, map actions, service-area leads.	GA4 events, call tracking, CRM
Enterprise / long cycle	Account engagement, assisted opportunity influence, sales enablement use, brand demand.	CRM, ABM platform, sales analytics

### Conversion Definition

Item	Definition
Primary conversion event #1	
Primary conversion event #2	
Primary conversion event #3	
Revenue or value proxy if direct revenue is unavailable	
Attribution caveats or known data gaps	

## 6. Data Access Matrix

Source	Access type	Minimum permission	Export alternative	Owner
Google Search Console	Read-only	Full or restricted property access	Page/query export by date range	
GA4 / analytics	Read-only	Viewer or analyst	Landing page, source/medium, events, conversions	
Sitemap / crawl	Public crawl or export	No login if public site	XML sitemap, CMS URL export, Screaming Frog crawl	
CRM	Read-only or export	Pipeline/revenue export only	Closed-won by source, URL, campaign, or landing page where available	
CMS	Export preferred	Content metadata only	URL, title, type, publish date, update date, author, status	
Warehouse / BI	Export preferred	Filtered extract	Landing page performance or revenue model	
Paid media	Read-only or export	Campaign/ad group/landing page data	Spend, clicks, conversions by landing page	

## 7. Channel Coverage Checklist

URL Ledger is channel-agnostic. The audit should capture every meaningful discovery and attribution surface, even when confidence varies by channel.

Channel	Input signal	What the ledger can measure
Organic search	GSC page/query, crawl, rank or SERP exports	Ranking, impressions, clicks, CTR, query coverage, indexation risk
AI answer surfaces	Manual samples, SERP screenshots, third-party monitoring, citation checks	Citation, omission, summarization, zero-click risk, competitor answer share
Paid search / paid social	Ad platform landing page exports	CAC, conversion efficiency, landing page quality, spend concentration
Social discovery	UTMs, referral source data, social analytics	Reusable content value, shareability, referral spikes
Email / lifecycle	Campaign click data, automation paths	Nurture contribution, retention, lifecycle-assisted URL value
Referral / backlinks	Backlink exports, referral traffic	Authority, trust transfer, partner-driven discovery
Direct / branded	Direct sessions, branded search, returning users	Brand recall, sales enablement, untracked demand
CRM / sales	Lead source, opportunity influence, closed-won	Pipeline influence, assisted conversion, sales-cycle value
Internal site search	Site search logs	User demand after arrival, content findability
AI agents / browser agents	DOM render checks, schema, accessibility, form flow	Machine readability, extractability, actionability, policy safety

## 8. 13 Structural Variables: Data Readiness View

#	Variable family	Primary inputs
1	Content Health	URL content, freshness, accuracy, depth, proof, usefulness
2	Content Decay	Historical traffic, rankings, conversions, freshness, engagement
3	Content Dilution	Inventory size, low-yield pages, duplicate patterns, topic sprawl
4	Cannibalization	Query overlap, internal competing pages, cluster conflict
5	Content Waste	Low-value pages, maintenance drag, crawl waste, redundant inventory
6	Content Investment	Production cost, update effort, owner, media/proof depth
7	Content ROI	Revenue, pipeline, conversion contribution, assisted value
8	Performance	Impressions, clicks, sessions, CTR, conversions, engagement
9	Intent / Journey Alignment	Query class, funnel role, page purpose, CTA fit
10	Messaging / Psychographic Fit	Audience mode, trust state, objections, decision friction
11	Authority / Entity Strength	Brand, author, topical authority, proof, links, citations
12	Technical / Indexation Structure	Canonical, redirect, crawl depth, schema, rendering, speed
13	AI Citation / Agent Readiness	Citation presence, extraction clarity, schema, DOM, forms, policy

### SCORING NOTE

Early audits can use directional confidence levels. The platform becomes more authoritative as history, actions, and outcomes write back into the ledger.

## 9. Security, Access, and Evidence Rules

- Read-only access is preferred for analytics, GSC, CRM, paid, and BI sources.
- Exports are acceptable when live access is not possible or when the client has strict security controls.
- No direct publishing, redirect changes, CMS changes, or code changes occur during intake.
- Every recommendation should be traceable to a source signal, confidence level, and expected measurement outcome.
- High-risk pages should be tagged as protected assets before any proof sprint or implementation step.
- Agent-facing use cases start as recommendation-only unless explicit action permissions are defined later.

Data type	Preferred handling	Retention / evidence note
Analytics exports	Store in client-specific audit folder	Used for baseline and scoring evidence
CRM / revenue exports	Filtered, minimum necessary fields	Used to model value; avoid unnecessary PII
Crawl files	Store with timestamp and crawl settings	Used for structural evidence and comparisons
Screenshots / SERP evidence	Capture with date and query context	Used for visibility, AI citation, and result layout proof
Recommendation logs	Versioned and tied to owner approvals	Used for audit trail and implementation handoff

## 10. Protected Asset Intake

Some URLs should not be changed casually, even when a score suggests an opportunity. These pages require explicit governance because they carry brand, legal, revenue, product, or trust sensitivity.

Asset type	Common examples	Default policy
Commercial trust pages	Pricing, demo, comparison, alternative pages	High approval
Legal / compliance pages	Terms, privacy, regulated claims	No-touch unless legal approved
Product truth pages	Feature, integration, security, API docs	Product owner approval
Brand-defining assets	About, category landing pages, flagship guides	Brand + strategy review
High-value organic pages	Top revenue, lead, or traffic URLs	Measurement plan required
Agent-sensitive flows	Forms, booking, checkout, login, calculators	Recommendation-only until policy exists

### Protected URL List

URL / cluster	Reason protected	Approver	Allowed action level

## 11. 45-Day Cadence and Client Commitments

Window	Workstream	Client-visible output
Day 0-5	Access + scope lock	Confirm data sources, permissions, domain boundaries, conversion model, and stakeholders.
Day 6-10	Inventory + baseline	Normalize URL inventory, map clusters, identify first-pass risk and value concentration.
Days 11-21	Structural scoring	Score URL assets across the 13 structural variable families and identify major decay drivers.
Days 22-30	Value-at-risk + repricing	Model portfolio leakage, confidence bands, recoverable value, and channel exposure.
Days 31-38	Backlog + proof path	Prioritize refresh, merge, retire, protect, expand, technical, and governance actions.
Days 39-45	Executive readout	Deliver findings, next-step operating model, ledger install path, and 90-day roadmap.

### MEETING RHYTHM

Recommended cadence: kickoff, Day 10 scope/baseline checkpoint, Day 25 backlog checkpoint, Day 45 executive readout. Async access questions can be handled between checkpoints.

## 12. Readiness Scorecard

Readiness area	Green	Yellow	Red
Scope clarity	Domains and exclusions clear	Some ambiguity by folder or market	No clear boundary
URL inventory	Sitemap/crawl/CMS export available	Partial inventory available	No reliable URL source
Search data	GSC available	Partial exports available	No page/query visibility
Analytics data	GA4/equivalent available	Limited landing page/event data	No analytics access
Revenue mapping	CRM/orders/revenue proxy available	Conversion events but weak value model	No agreed conversion definition
Stakeholders	Data, technical, content, revenue owners named	Some owners missing	No accountable owner
Governance	Protected assets and approvals known	Some sensitivity known	No policy clarity

## 13. Intake Questionnaire

Area	Question
Business context	What triggered interest now? Core update volatility, traffic decline, revenue drift, content waste, AI visibility concern, migration, redesign, budget scrutiny, or agent governance?
Portfolio scale	Approximately how many URLs, domains, subdomains, languages, or markets are in scope?
Revenue dependence	How much of pipeline, ecommerce revenue, lead flow, or customer acquisition depends on the website?
Known decay	Which clusters, page types, or channels appear to be declining or underperforming?
Known waste	Where do you suspect redundant, outdated, low-yield, or unmaintained content exists?
Conversion model	Which events or revenue outcomes should define URL yield?
Attribution gaps	Where do current dashboards fail to connect content, traffic, conversion, and revenue?
AI / agent concern	Are AI answers, zero-click exposure, AI citations, browser agents, or internal AI workflows part of the concern?
Protected assets	Which pages should not be edited, merged, redirected, or touched without approval?
Success definition	What would make the audit worth doing: recovered revenue, backlog clarity, governance install, budget proof, agent readiness, or platform pilot?

## 14. Handoff Checklist Before Audit Begins

Checklist item	Definition of done
Scope locked	Domains, subdomains, folders, languages, and exclusions confirmed
Stakeholders named	Executive sponsor, data owner, technical owner, content owner, revenue owner
Access path confirmed	Read-only access or export plan for all required sources
Conversion model agreed	Primary events and value proxies defined
URL inventory available	Sitemap, crawl permission, CMS export, or existing URL list ready
Protected assets flagged	No-touch and high-approval pages identified
Security handling approved	Data transfer, storage, retention, and evidence expectations agreed
Meeting cadence booked	Kickoff, Day 10, Day 25, and Day 45 checkpoints scheduled
Readout audience defined	CFO, CMO, Growth, RevOps, SEO, content, engineering, or investor stakeholders

## 15. Audit-to-Ledger Conversion Path

The intake is not just pre-work. It establishes the first layer of the system of record. The same inputs used to run the audit become the starting data model for the ongoing URL Ledger.

Audit input	Ledger object created	Future use
URL inventory	Canonical URL registry	Asset identity, dedupe, lineage, ownership
GSC + analytics	Performance history	Decay curves, visibility, traffic and conversion patterns
CRM / revenue model	Yield model	Value-at-risk, ROI, recovery priority
Crawl + technical signals	Structural risk record	Indexation, canonical, redirect, crawl depth, rendering risk
Protected asset list	Policy object	Approval gates, no-touch rules, agent action limits
Backlog recommendations	Action queue	Refresh, merge, retire, protect, expand, technical fix, proof sprint
Readout decisions	Governance history	What was approved, deferred, rejected, or escalated

### CONVERSION THESIS

The audit is the first governed interpretation. The ledger turns that interpretation into ongoing operational truth.

## 16. Final Intake Notes

- The strongest audits start with clear scope, a practical conversion model, and enough access to connect URL behavior with business value.
- Perfect attribution is not required. Directional evidence is enough to expose decay, waste, and repricing opportunities.
- The client should understand that the deliverable is not only a report. It is a ranked operating backlog and the first install path for the ledger.
- AI Search, AI agents, and zero-click behavior should be captured as emerging discovery surfaces, not as the entire category.
- The product category remains website asset intelligence: a system of record for URL asset value across channels, teams, tools, and future agents.

**Next step: confirm scope, conversion definition, access path, and the Day 0 kickoff date.**