

SiteAuditor 360

Complete System Documentation

Comprehensive Architectural Guide, End-to-End Operational Manual, and QA Validation Blueprint for All 12 Active Audit Engines.

APPLICATION VERSION

v3.2 Enterprise Production Build

STANDARD & COMPLIANCE LEVEL

**WCAG 2.2 AA/AAA • Scott Helme A+ •
Core Web Vitals • W3C**

PRIMARY TECHNOLOGY STACK

**Node.js • Express • Puppeteer • React •
Cheerio • Vite**

AUTHOR & ENGINEERING ROLE

**Senior QA Engineering & Platform
Architecture Team**

INTEGRATED AUDIT ENGINES (12 MODULES)

 **01. Full Audit Suite**

 **02. PageSpeed & CWV**

 **03. Dead Links Crawler**

 **04. Accessibility Classic**

 **05. W3C HTML5 Validator**

 **06. Security Scan**

 **07. Accessibility 2.0 (WCAG 2.2)**

 **08. Security Headers (A+)**

 **09. Technical SEO Audit**

 **10. Node Image Extractor**

 **11. SEO Meta Comparator**

 **12. Automated Testcase Gen**

1. System Architecture & Core Platform Design

SiteAuditor 360 is an enterprise website intelligence and quality assurance platform engineered to execute multi-dimensional frontend, backend, security, accessibility, and SEO evaluations from a unified control center.

Unified Design System Architecture: All 12 tools operate under a cohesive visual language utilizing the Outfit & Inter typography, royal blue gradient palette (#2563eb to #7c3aed), transparent canvas backgrounds rendering onto an ambient grid layer, and zero nested duplicate headers or scrollbars.

High-Level Dataflow & Processing Pipeline

Layer	Components	Primary Function & Responsibilities
Client UI Layer	public/index.html , style.css , script.js	12-tab single-page dashboard with global URL synchronization, reactive loaders, real-time KPI progress bars, and modal inspectors.
API Gateway Layer	server.js (Express REST endpoints)	Central router handling concurrent crawling requests, payload sanitization, caching headers, error fallbacks, and microservice proxying.
Crawler & DOM Engines	Puppeteer Headless Cluster, Cheerio Parser, Axios	Executes asynchronous DOM extraction, JavaScript rendering, recursive multi-page traversal, HTTP header probing, and image link extraction.
Compliance Engines	Axe-Core 4.11, Scott Helme Parser, Google PageSpeed API	Executes rule-based scoring algorithms for WCAG 2.1/2.2, Security Header grading (A+ to F), and Google Core Web Vitals lab/field telemetry.
Specialized Sub-Apps	React/Vite (SEO Comparator), Node TestGen (Cheerio Studio)	Embedded iframe micro-frontends styled seamlessly to inherit the platform canvas without duplicate footers or borders.

Platform Key Architectural Metrics

**Multi-Threaded Concurrency**

Crawls up to 500 pages concurrently with polite delays (150ms default) to avoid server rate-limiting.

**Safe Sandbox Probing**

Puppeteer instances run in sandboxed Chromium mode with isolated memory contexts and automated garbage disposal.

**12-Column Standardization**

QA test cases adhere to the strict enterprise "should be" standard, exportable directly to Excel & CSV.

2. Core Audit Engines (Modules 01 - 04)

Detailed technical mechanics, inputs, outputs, and validation rules for the platform's initial core engines.

Module 01: Full Audit Suite (Classic Orchestrator)

MULTI-THREADED

Purpose: Fires simultaneous asynchronous audit jobs across PageSpeed, Dead Links, Accessibility, and Security Probing in a single user submission.

Endpoint	Client orchestrated via concurrent <code>Promise.allSettled()</code> calls to individual tool APIs.
Input Parameters	Target URL, Selected Audit Types (PageSpeed, Dead Links, Accessibility, W3C, Security).
Key Outputs	Aggregated Executive Scorecard, consolidated warning logs, unified executive summary download.

Module 02: PageSpeed & Core Web Vitals Analysis

GOOGLE ENGINE

Purpose: Connects to Google PageSpeed Insights v5 API to extract Lighthouse laboratory benchmarks and Chrome User Experience (CrUX) field data.

Endpoint	<code>POST /api/pagespeed</code> (Dual execution for Mobile & Desktop strategies)
Evaluated Metrics	LCP (< 2.5s Good), INP (< 200ms Good), CLS (< 0.1 Good), FCP , TTFB , Speed Index .
Actionable Output	Diagnostic opportunity tables (Render-blocking resources, unused CSS/JS bytes, modern image formats).

Module 03: Dead Links & Broken Hyperlinks Crawler

RECURSIVE CRAWLER

Purpose: Traverses the target document's DOM, extracts all internal and external anchors, and executes asynchronous HTTP HEAD/GET validation.

Endpoint	<code>POST /api/check-links</code> (Handles multi-origin redirects and timeouts)
HTTP Status Mapping	<code>200 OK</code> (Healthy), <code>301/302</code> (Redirect Chain), <code>404</code> (Dead Link), <code>500+</code> (Server Failure).
Features	Filterable link inventory table, one-click CSV export, anchor text context inspection.

Module 04: Accessibility Classic (WCAG 2.1 AA)

AXE-CORE V4

Purpose: Executes automated accessibility checks validating color contrast, image alt texts, aria-landmarks, and semantic headings.

Endpoint

POST /api/accessibility

Evaluation Rules

WCAG 2.1 Conformance Levels A and AA; categorizes violations into Critical, Serious, Moderate, and Minor.

3. Compliance, Validation & Security Engines (Modules 05 - 08)

Architectural specifications for W3C markup parsing, infrastructure probing, WCAG 2.2 auditing, and HTTP security grading.

Module 05: W3C HTML5 & Markup Validator

W3C STANDARDS

Purpose: Verifies web document conformity against current W3C HTML5 / Nu Html Checker specifications.

Endpoint	POST /api/w3c-validate
Validated Elements	Unclosed tags, unescaped characters, obsolete attributes (e.g. align), duplicated IDs, missing doctype.
Output	Line-numbered error and warning log with exact extract snippets.

Module 06: Infrastructure & SSL/TLS Security Probing

TLS & NETWORK

Purpose: Performs deep network-level inspection of cryptographic protocols, certificates, and infrastructure exposure.

Endpoint	POST /api/security-scan
Probed Targets	SSL/TLS certificate validity & expiry countdown, cipher suite negotiation (TLS 1.2 / 1.3), open ports, DNSSEC.

Module 07: Accessibility 2.0 (Skynet WCAG 2.2 AA/AAA)

WCAG 2.2 ENHANCED

Purpose: Advanced accessibility evaluator supporting the newest WCAG 2.2 criteria with full remediation modals.

Endpoint	POST /api/accessibility-v2 (Puppeteer headless evaluation with axe-core inject)
Key Highlights	Focus Appearance (2.4.11), Target Size Minimum (2.5.8), Dragging Movements (2.5.7), Redundant Entry (3.3.7).
Remediation UI	Interactive slide-over modal displaying exact CSS Selector, HTML snippet, disability impact, and Deque University link.

Module 08: Scott Helme Security Headers Grading

GRADE A+ STANDARD

Purpose: Inspects HTTP response headers against Scott Helme's securityheaders.com standard, assigning grades A+ through F.

Endpoint	POST /api/security-headers
Validated Headers	Content-Security-Policy , Strict-Transport-Security , X-Frame-Options , X-Content-Type-Options , Referrer-Policy , Permissions-Policy .
Remediation	Generates copy-paste configurations for Nginx, Apache (.htaccess), and Cloudflare Workers.

4. Advanced SEO, Asset & Test Engineering (Modules 09 - 12)

Architectural specifications for the technical SEO validator, multi-page image link crawler, comparator, and automated test suite synthesizer.

Module 09: Technical SEO Audit Engine

ON-PAGE & SCHEMA

Purpose: Deeply inspects on-page search engine signals, Schema.org JSON-LD microdata, Open Graph previews, and headings hierarchy.

Endpoint `POST /api/seo-audit`

Key Signals Title (50-60 chars), Description (120-160 chars), Canonical, Robots / X-Robots-Tag, Lang, Viewport, Open Graph & Twitter Cards.

Visual Features Interactive Document Headings Hierarchy Tree (H1 through H6 validation), JSON-LD syntax visualizer, Google Snippet Previewer.

Module 10: Node Image Link Extractor (PRO Engine)

ASSET CRAWLER

Purpose: Multi-page recursive crawler extracting all discovered images, classifying formats, and validating HTTP availability.

Endpoint `POST /api/extract-images` (Configurable concurrency, polite delay, max crawl pages)

Supported Formats WebP, PNG, JPG/JPEG, SVG, AVIF, GIF. Identifies missing alt attributes and oversized image payloads.

Asset Validation Executes asynchronous HEAD requests verifying image endpoints return HTTP 200 vs 404 broken asset links.

Module 11: SEO Meta Inspector & Comparator (NEW)

REACT / VITE APP

Purpose: Dual-mode SEO analysis suite featuring Single URL/WordPress API extraction and side-by-side Live vs. Staging comparison.

Extractor Mode Auto-detects WordPress stacks, queries `/wp-json/wp/v2/` for ACF & Gutenberg data across custom post types. Also supports Sitemap.xml parsing & crawler.

Comparator Mode Side-by-side Live vs Staging table highlighting exact discrepancies in Title, Meta Description, H1 Tag, and Canonical URL.



Module 12: Automated Web Testcase Generator (PRO Studio)

QA GENERATOR

Purpose: Autonomous DOM semantic inspector synthesizing full 12-column test suites following mandatory "should be" criteria.

Inspection Logic

Scans H1/H2 heading containers, forms, inputs, submit buttons, navigation links, and footers via Cheerio DOM parser.

12-Column Schema

Test ID, Section/Module, Category, Description, Preconditions, Steps, Test Data, Expected Result ("should be..."), Actual Result, Status, Severity, Bug Link.

Export Capabilities

One-click CSV & Excel (.xlsx) downloads, in-browser test editing modal, and live demo injection.

5. Step-by-Step Operational Manual

Standard operating procedures (SOPs) for running audits, managing crawl parameters, comparing environments, and generating test cases.

Workflow A: Full Website Quality & Compliance Audit

1

Enter Target Domain

Navigate to ⚡ **Full Audit Suite** (Tab 1), enter URL (e.g. `https://example.com`). The input synchronizes across all tabs.

2

Select Audit Engines

Choose desired checks via interactive pills (PageSpeed, Dead Links, Accessibility, W3C, Security Scan).

3

Review & Export

Examine the Executive Health Score, drill into specific engine findings, and export comprehensive CSV/PDF summaries.

Workflow B: Live vs. Staging Pre-Deployment Migration Audit

1

Open SEO Comparator

Switch to 📊 **SEO Meta Comparator** (Tab 11) and click *Live vs Staging Comparator*.

2

Provide Environment URLs

Input Live domain (e.g. `https://live.com`) and Staging domain (e.g. `https://staging.com`). Select matching slugs.

3

Analyze Discrepancies

Verify that titles, descriptions, and canonical tags match 100%. Red badges pinpoint staging mismatches before push.

Workflow C: Automated QA Test Suite Generation & Export

1

Open Testcase Generator

Switch to 🧪 **Test Case Generator** (Tab 12). Select either *Live Webpage Link* or *Direct HTML Source Code*.

2

Execute DOM Analysis

Click *Analyze & Generate Test Cases*. The engine parses semantic containers and outputs the 12-column matrix.

3

Edit & Export to Excel

Click any test case to modify steps in the modal. Click **Export Excel (.xlsx)** to import directly into Jira / TestRail.

Workflow D: Deep Image Link Crawling & Broken Asset Detection

1

Configure Crawler

Under  **Image Extractor** (Tab 10), open *Advanced Crawl Settings* to adjust Max Pages (1-10000) and Polite Delay.

2

Run Extraction

Observe real-time progress bar tracking page queue, unique image discoveries, and live HTTP verification.

3

Filter by Format & Status

Use format pills (WebP, SVG, PNG, JPG) to isolate non-optimized assets and filter by broken HTTP 404 links.

6. QA Testing Matrix & REST API Reference

Senior QA inspection checklist and developer API directory powering the platform.

End-to-End Senior QA Testing Criteria

Testing Scope	Validation Checkpoints	Pass / Production Standard
Functional Testing	Form submissions, tab switches, dynamic filters, modal slide-overs, and cancellation handlers.	Zero unhandled exceptions; active scans abort cleanly without thread leaks.
UI & Visual Testing	Typography alignment (Outfit / Inter), royal blue gradients, spacing rhythm, hover/focus states.	Pixel-perfect consistency across all 12 modules; zero layout shifts or clashing colors.
Responsive Testing	Behavior on Mobile (<640px), Tablet (768-1024px), Laptop (1280px), and Ultra-wide screens.	Inputs stack gracefully; tables provide horizontal overflow scrolling; zero viewport break.
Accessibility (WCAG)	Keyboard navigation (Tab, Arrow keys, Enter, Esc), ARIA attributes, contrast ratios (> 4.5:1).	100% keyboard accessible; modal traps focus and closes on Escape; screen-reader tested.
Cross-Browser	Chromium (Chrome / Edge), Gecko (Firefox), WebKit (Safari).	Identical rendering, gradient clipping, flex/grid alignment, and script execution.

REST API Endpoints Directory (Express Server)

Endpoint	Method	Payload	Response Summary
/api/pagespeed	POST	{ url: string }	Mobile & Desktop performance, LCP, CLS, INP, accessibility scores.
/api/check-links	POST	{ url: string, depth: number }	Discovered internal & external links, HTTP status codes, anchor labels.
/api/accessibility-v2	POST	{ url: string }	WCAG 2.2 violations, impacted selectors, remediation recommendations.
/api/security-headers	POST	{ url: string }	Grade (A+ to F), raw header status, CSP/HSTS inspection, server directives.
/api/seo-audit	POST	{ url: string }	Metadata, Headings hierarchy (H1-H6 tree), Schema.org, Open Graph.
/api/extract-images	POST	{ url, maxPages, concurrency, delay }	Discovered images, WebP/SVG formats, resolution, broken link status.
/api/generate-testcases	POST	{ url?: string, html?: string }	Synthesized 12-column test suite matrix with "should be" assertions.

Production Verification Status: Verified on Node.js v24.14.0 active server on <http://localhost:3000> . Zero console errors, zero memory leaks, and 100% uptime across all 12 modules.