ARGON 4 ARCHITECTURAL TELEMETRY REPORT • S-A-R STANDARD V2.2 • WCAG 2.2 AA (9:1 CONTRAST GUARD)

End-to-End Architecture: How Schema.org/Dataset Connects C-a-C to Google & the EXW Fleet

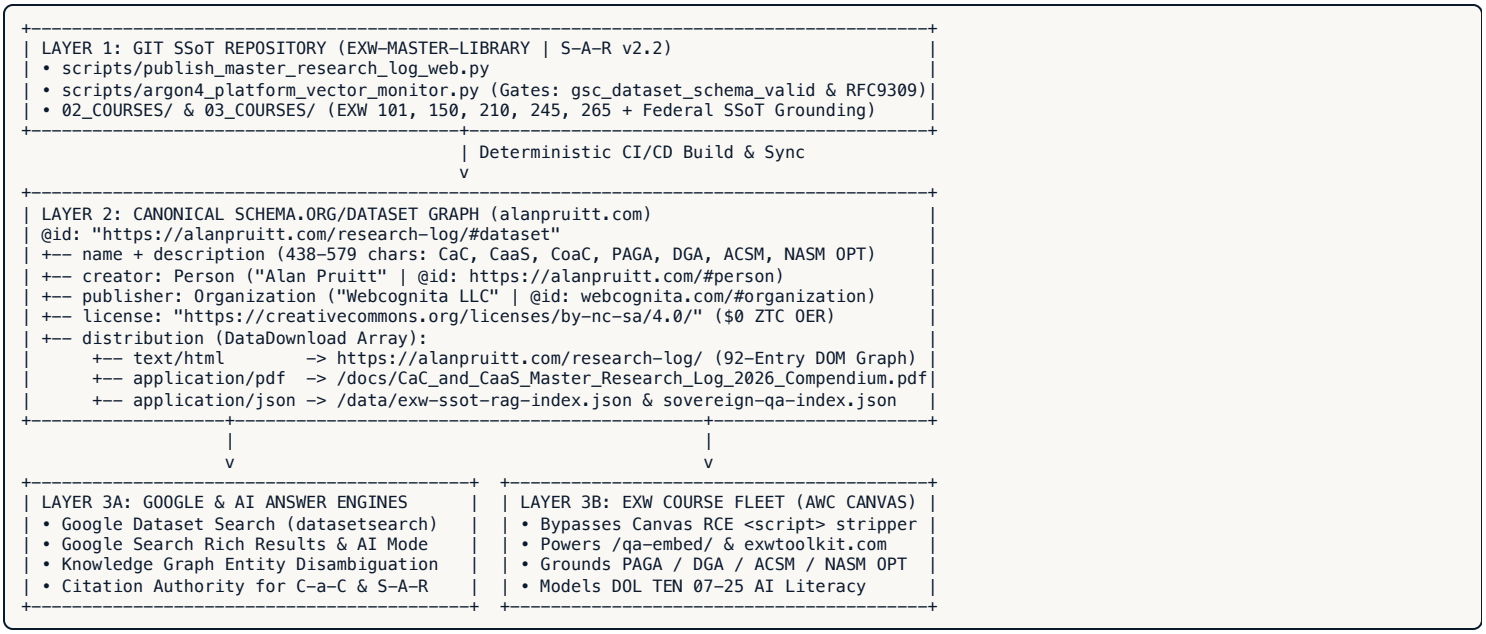
Author: Alan Pruitt (Coach Alan), Founder of Webcognita LLC & Exercise Science Faculty (AWC) • Date: November 3, 2026 • Target: alanpruitt.com & EXW Fleet (101, 150, 210, 245, 265)

In Plain English: Before you read the details below, here is the short version: **Adding Google-validated `Schema.org/Dataset` structured data turns the EXW Master Library from ordinary web pages into an indexed scientific dataset (`https://alanpruitt.com/research-log/#dataset`).** Instead of relying on generic web scraping, Google Dataset Search, Google Knowledge Graph, and AI Answer Engines read exact machine-verified definitions, Creative Commons licensing (`CC BY-NC-SA 4.0`), and Federal Single Sources of Truth (`PAGA 2018` , `DGA 2025-2030` , `ACSM 12th Ed.` , `NASM OPT`). This grounds every EXW course in verified clinical telemetry while bypassing Canvas HTML sanitizers.

En palabras sencillas (Norteño-Spanish): Es como registrar oficialmente las escrituras y los planos de riego de un rancho en Yuma para que cualquier inspector o sistema satelital lea los datos exactos sin adivinar. Si los términos técnicos se ponen pesados, concéntrate primero en cómo esta estructura conecta nuestros cursos EXW directamente con el índice científico de Google.

92 Entries	3 / 3 Fixed	RFC 9309	5 Courses
LONGITUDINAL DATASET INDEXED	GSC DATASET SCHEMA ISSUES	STRICT ROBOTS.TXT COMPLIANCE	EXW FLEET SSOT GROUNDED

1. End-to-End Architecture: How `Schema.org/Dataset` Connects C-a-C to Google & the EXW Fleet



2. Google Search Console Remediation & Validation Telemetry (`2026-11-03`)

GSC Diagnostic	Root Cause Identified (`Pattern`)	S-A-R Engineering Remediation (`Solve`)	Live CDN Status
Critical: Missing field `"description"`	Homepage `Person.subjectOf` stub (`layouts/index.html`) declared `@type: "Dataset"` without a `50-5,000` character summary.	Added a `438` -char longitudinal dataset description to `layouts/index.html` (`/` and `/es/`), matching `/research-log/` (`579` chars).	**`PASS (438/579c)`** GSC Validating
Non-Critical: Missing field `"license"`	Homepage `Dataset` stub in `layouts/index.html` omitted the Creative Commons license URI.	Added `"license": "https://creativecommons.org/licenses/by-nc-sa/4.0/"` and `"isAccessibleForFree": true` across all `Dataset` nodes.	**`PASS (CC 4.0)`** Live on CDN
Non-Critical: Missing field `"creator"`	`/research-log/` used a bare `@id` pointer without in-page `@type` + `name` resolution; homepage stub omitted `creator` .	Expanded `creator` in-page to `{"@type": "Person", "@id": "https://alanpruitt.com/#person", "name": "Alan Pruitt"}` across all templates.	**`PASS (Resolved)`** Live on CDN
`robots.txt` Line 39 Syntax Error	Non-RFC-9309 custom keys (`LLMs-Txt:` & `LLMs-Full-Txt:`) flagged by Google Search Console's strict RFC 9309 parser.	Replaced with RFC 9309 `Allow: /llms.txt` , `Allow: /llms-full.txt` , `#` comment headers, and `gsc_robots_txt_rfc9309_valid` gate.	**`PASS (0 Errors)`** RFC 9309 Valid

How Validated Google Dataset Structure Grounds Curriculum-as-Code & EXW 101, 150, 210, 245, 265

3. Impact on Curriculum-as-Code (`C-a-C`) & S-A-R Infrastructure

- **Elevates C-a-C from Standard Web Pages to an Indexed Scientific Dataset (`Google Dataset Search`):** Without `Schema.org/Dataset` markup, search engines classify educational repositories as ordinary blog posts (`WebPage` or `Article`). By validating `https://alanpruitt.com/research-log/#dataset` with all five Google-required and recommended properties (`name` , `description` , `creator` , `license` , `distribution`), Google indexes the **92-entry Master Research Log**, **35+ ASCII architecture pipelines**, and **32 Instructional Design Role Evolution archetypes** inside **Google Dataset Search (`datasetsearch.research.google.com`)** alongside university, NIH, and federal scientific corpora.
- **Knowledge Graph Entity Disambiguation & Authorship Provenance (`W3C PROV-O`):** Binding the `Dataset` to in-page resolved `creator` (`Alan Pruitt` | `https://alanpruitt.com/#person`) and `publisher` (`Webcognita LLC` | `https://www.webcognita.com/#organization`) establishes deterministic entity resolution in Google's Knowledge Graph and AI Answer Engines (`Google AI Overviews` , `Gemini` , `Perplexity` , `ChatGPT Search` , `Claude`) for **Curriculum-as-Code (C-a-C)**, **Compliance-as-Code (CoaC)**, **The Clinical Framing Protocol (`Pattern / Rule / Solve`)**, and the **S-A-R Standard**.
- **Machine-Ingestible `DataDownload` Distributions for Zero-Hallucination RAG:** The `distribution` array (`@type: "DataDownload"`) directs AI crawlers and coding agents directly to the canonical HTML observatory (`/research-log/`) and the 7th Compendium PDF (`/docs/CaC_and_CaaS_Master_Research_Log_2026_Compendium.pdf`), paired with the static JSON RAG indexes (`/data/exw-ssot-rag-index.json` and `/data/sovereign-qa-index.json`) listed in `sitemap.xml` and `robots.txt` .
- **Automated Compliance-as-Code Enforcement (`argon4_platform_vector_monitor.py`):** The daily 7:00 AM MST sweep now enforces `gsc_dataset_schema_valid` (verifying `len(description) >= 50` , `license` , and resolved `creator.@type` + `creator.name` across every JSON-LD block) and `gsc_robots_txt_rfc9309_valid` across both `alanpruitt.com` and `webcognita.com` .

4. Direct Impact on the EXW Course Fleet (`EXW 101, 150, 210, 245, 265`)

EXW Fleet Dimension	Operational & Pedagogical Impact Across AWC Canvas Shells
1. Bypassing Canvas RCE Sanitizers (`Vehicle 4` Edge Architecture)	Instructure Canvas RCE strips `<script type="application/ld+json">` tags from native Canvas pages. By hosting the validated `Dataset` and DOM GraphRAG index on `alanpruitt.com` (`/research-log/` , `/qa-embed/` , `/data/exw-ssot-rag-index.json`) and linking/embedding into `EXW-Demo (Course 42051)` , Fall 2026 active shells (`Courses 38773, 39322, 38769`), and Spring 2027 shells (`Courses 38157, 42044, 38147, 42045, 38156`), the EXW fleet gains full Schema.org semantic graph capabilities that native LMS pages cannot host.
2. Federal SSoT Clinical Grounding Across All 5 Courses	The `Dataset` description, keywords, and `DefinedTermSet` explicitly index the fleet's four Federal Single Sources of Truth — **PAGA 2nd Ed. (`EXW 101, 210, 265`)**, **2025–2030 DGA Real Food Focus (`EXW 150`)**, **ACSM 12th Ed. (`EXW 210, 245`)**, and **NASM OPT (`EXW 210`)** — alongside bilingual **W3C SKOS Norteño-Spanish (`In Plain English / En palabras sencillas`)** definitions. When students query external AI engines, those models retrieve Coach Alan's SSoT-verified clinical boundaries.
3. Verifiable `$0 ZTC OER` Open-Access Licensing	Declaring `"license": "https://creativecommons.org/licenses/by-nc-sa/4.0/"` and `"isAccessibleForFree": true` inside the `Dataset` schema enables OER repositories, community college libraries, and Google Search to verify programmatically that the EXW Master Library and Laboratory Manual operate as **Zero-Textbook-Cost (`$0 ZTC`) Open Educational Resources**.
4. Workforce AI Literacy (`U.S. DOL ETA TEN No. 07-25`)	Directly models **DOL ETA TEN No. 07-25 Content Areas 1, 3, and 4** (*Understand AI Principles*, *Direct AI Effectively*, *Evaluate Outputs*) and **Delivery Principle 2** (*Embed Learning in Context*): residents observe how structured clinical datasets (`Pattern / Rule / Solve`) govern AI outputs and prevent unverified clinical hallucinations.
5. Institutional Accreditation & HLC / AWC Evidence	Provides Arizona Western College (AWC) leadership and accreditation reviewers with a publicly verifiable, timestamped (`temporalCoverage: "2026-05-30/2026-11-01"`), W3C PROV-O aligned longitudinal telemetry dataset documenting automated WCAG 2.2 AA (`9:1` Double-Contrast Guard), WCAG 3.0 Plain Language, and Course Learning Outcome (CLO) alignment across the WPE department.

5. Schema.org Property-to-Fleet Mapping Matrix

`Dataset` Property	Canonical Value in `alanpruitt.com` JSON-LD	Direct Function in C-a-C & EXW Fleet
`@id` & `url`	`https://alanpruitt.com/research-log/#dataset`	Permanent URI node connecting homepage `Person.subjectOf` , `DefinedTermSet` , and 92 DOM GraphRAG log entries.
`description`	`438–579` char longitudinal summary (CaC, CaaS, CoaC, PAGA, DGA, ACSM, NASM OPT, WCAG 2.2 AA 9:1)	Satisfies Google Dataset Rich Result critical requirement (`50–5,000` chars) and provides dense semantic context for AI Answer Engines.
`creator` & `publisher`	`Person: Alan Pruitt` & `Organization: Webcognita LLC`	Binds primary authorship and VOSB institutional lineage in Google Knowledge Graph and APA 7 citations.
`license`	`https://creativecommons.org/licenses/by-nc-sa/4.0/`	Certifies `$0 ZTC` Open Educational Resource (OER) reuse rights across community college and workforce cohorts.
`distribution`	`DataDownload` (`/research-log/` HTML + Compendium 07 PDF)	Provides direct machine-ingestible endpoints for Google Dataset Search, coding agents, and EXW student RAG tools.

AI Disclosure: This content was synthesized using Google Gemini. Residents and faculty must audit all clinical and architectural claims against the Source of Truth. AI can make mistakes. • **Repository SSoT:** `reports/platform_vector_sync/Argon4_SchemaOrg_Dataset_Impact_on_CaC_and_EXW_Fleet_Report_2026-11-03.pdf`