

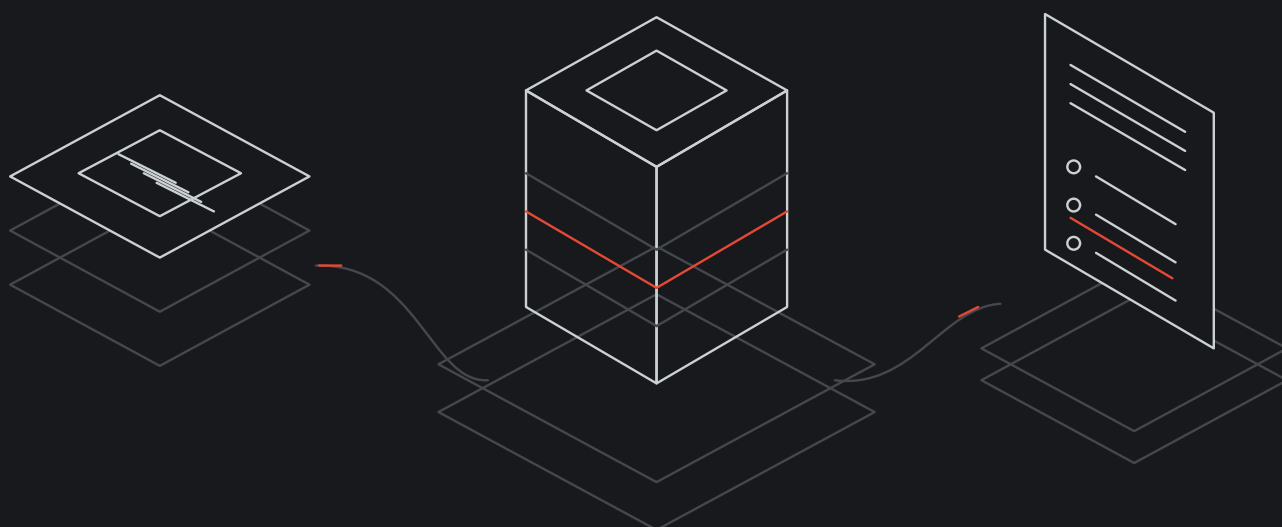
MACHINE EXPERIENCE AUDIT

MX Audit

Specimen report

A publishing portfolio

Content access, structure and discovery



SOURCE MATERIAL

Five real audit reports

PRESENTATION EDITION

11 October 2026

Composite specimen. Publisher identities and URLs withheld.

Recorded measurements preserved; evidence limits stated.

READING THIS SPECIMEN

Inside the report

Executive Summary	3
Scope and source material	4
Recorded MX measurements	5
Prioritised action plan	6
Method and evidence limits	7
Page-level quality and readability	8
Response failures need URL-level investigation	19
Served and rendered content need comparison	20
Structured-data coverage varies by template	21
Discovery inventory needs reconciliation	22
Page regions need explicit structure	23
Simulated crawler responses need policy context	24
One reached page cannot represent a site	25
Size warnings are heuristic signals	26
Delivery evidence register	27
Entity and template evidence	28
Source reconciliation register	29
Supporting Technical Findings	30
Security and response policies	31
Response timing observations	32
Optional conventions and experiments	33
Re-test method and acceptance criteria	34
Source traceability and attribution	35
Technical references	36

Pages 3-7 provide the decision overview. The findings that follow distinguish recorded observations from the extra evidence needed to verify their impact.

01 / DECISION OVERVIEW

Executive Summary

The immediate question is whether the intended content arrives, with enough structure for software to use it. The source reports flag delivery differences, response failures and uneven structured-data coverage. Investigate those first; optional MX markers are secondary.

What deserves attention

MX focus	What the sources record	Decision
Content delivery	Served-versus-rendered flags in four full reports	Inspect exact content differences before choosing a rendering fix
Response reliability	19 non-200 requests in A; one in B	Recover response bodies and re-test affected URLs
Entity descriptions	JSON-LD on 1/81 pages in A and 1/40 in B	Review representative content templates and intended consumers
Discovery and scope	Unreconciled sitemap inventories; C covers one page	Confirm coverage before making site-wide claims

This composite uses five separate, real source reports. Publisher A-E are anonymous source labels, with each sample kept separate. Each recommendation names the source evidence it depends on.

Accessibility, performance and security are supporting technical findings. A high aggregate pipeline score is not evidence that its failed checks can be ignored.

02 / DECISION OVERVIEW

Scope and source material

Five independent samples, with different coverage. Their recorded counts remain separate.

Source	HTML pages assessed	Source report	Coverage limit
Publisher A	81	84-page report; 7 Oct 2026	No readable sitemap recorded; whole-site share unknown
Publisher B	40	68-page report; 7 Oct 2026	Reported sitemap and expanded-index counts do not reconcile
Publisher C	1	4-page summary; date not stated	Only one page reached; sitemap lists 36149 URLs
Publisher D	52	70-page report; 7 Oct 2026	Bounded sitemap read; later parts may be unrepresented
Publisher E	100	80-page report; 7 Oct 2026	Sample drawn from a reported 30421-URL sitemap

HTML and PDF are separate scopes

The four full reports record no discovered PDFs in their samples. C does not provide a PDF-discovery result. No PDF content assessment is evidenced by these five reports. Under the stated service scope, the first PDF discovered is reviewed; additional PDF reviews are separately scoped.

The evidence references throughout use Source A-E and original PDF page numbers. Real publisher names, domains, article titles and paths are withheld in this public specimen. Counts describe the source samples, not the entire organisations.

03 / DECISION OVERVIEW

Recorded MX measurements

Use the measurements to guide investigation. They are original scanner outputs, not a universal certification scale.

Metric (out of 100)	A	B	C*	D	E
Served-HTML structure	42	27	34	20	71
Discovery readiness	10	25	40	32	40
Structured data quality	1	2	62	38	73
MX stack completeness	36	59	69	51	71
Agent readability	34	50	Not stated	68	70
Pipeline survivability	78	67	58	67	79

Sources: A pp4/20/45; B pp4/12/36; D pp4/13/37; E pp4/20/44; C p4. *C is a single-page result and supplies no detailed pipeline or schema register.

What these scores cannot establish

The source rubrics include optional discovery and MX conventions. Their exact weights and several raw sidecars are unavailable here. We preserve the outputs without inventing new weights, peer comparisons or an overall portfolio score.

A and B have displayed SDQ component totals that do not equal the reported aggregate; D also has a discrepancy. Rounding or aggregation may be involved, but the calculation was not supplied. The reconciliation register later states both values.

Served-HTML structure is not the same measure as the semantic-container score. Pipeline and served-versus-rendered flags need content-level review, even when a headline score is high.

04 / DECISION OVERVIEW

Prioritised action plan

Work from access and evidence to structure and metadata.

Order	Action	Completion evidence
1	Recover raw responses, paired HTML, URL sidecars and scoring calculations	Affected URL list and actual bodies available; unresolved counts explained
2	Re-test failed responses and intended article delivery	Valid, error and browser views compared under the same access conditions
3	Review representative article, author, listing and recipe templates	Primary content, headings, source metadata and relevant entities match
4	Reconcile canonical sitemap inventory and bot policy	Intended URLs and permissions checked; duplicates and exclusions justified
Parallel	Review accessibility, security and performance separately	Contextual specialist checks and recorded before/after results

Keep access intentional

Machine readability does not require allowing every crawler or authorising training use. Compare bot-specific rules with the publisher's intended policy before changing permissions. Make the evidence and the target consumer part of the agreed scope.

These are the specimen's recommended action order, not confidence labels or a replacement for each source's severity classification. Effort and remediation ownership should be agreed after validating the affected templates.

05 / DECISION OVERVIEW

Method and evidence limits

What is available	What it supports
Five supplied report PDFs	Recorded values, named tests, original tables and narrative context
Original report page references and file digests	Traceability to a specific supplied source edition
Raw HTML / browser snapshots not supplied	Actual missing text and extraction accuracy remain to be verified
Several referenced CSV/JSON sidecars not supplied	Full affected URL lists and score calculations cannot be independently reconstructed
No new crawl performed for this edition	These are historical sample observations, not current website verdicts

Severity and confidence

A severe impact can still have uncertain evidence. The source system has not supplied consistent Confirmed / Observed / Potential determinations, so this edition does not assign that taxonomy or imply a re-test happened. Each finding instead states the observation, evidence, possible impact, recommendation and verification needed.

Standards and source conventions

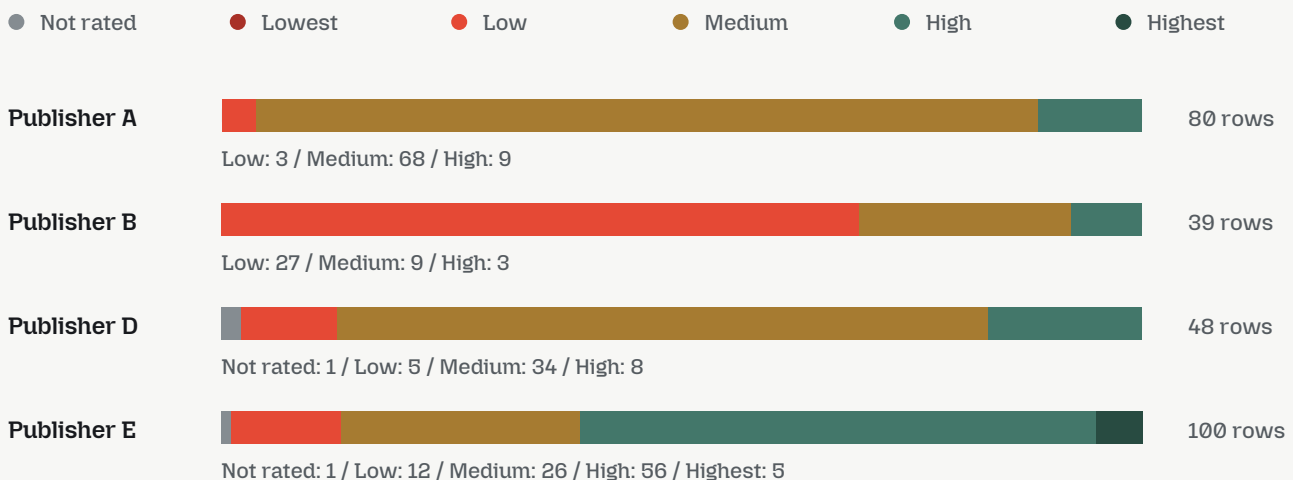
HTML delivery, HTTP semantics, robots rules and accurate structured data are the fundamentals. MX-specific maturity labels and metadata associated with The Gathering are source-framework conventions; they are not Tityra-owned standards. Tityra is the service brand of Digital Domain Technologies Ltd, trading as CogNovaMX. The full reports name Tom Cranstoun and the Web Audit Suite.

06 / PAGE MAPS

Page-level quality and readability

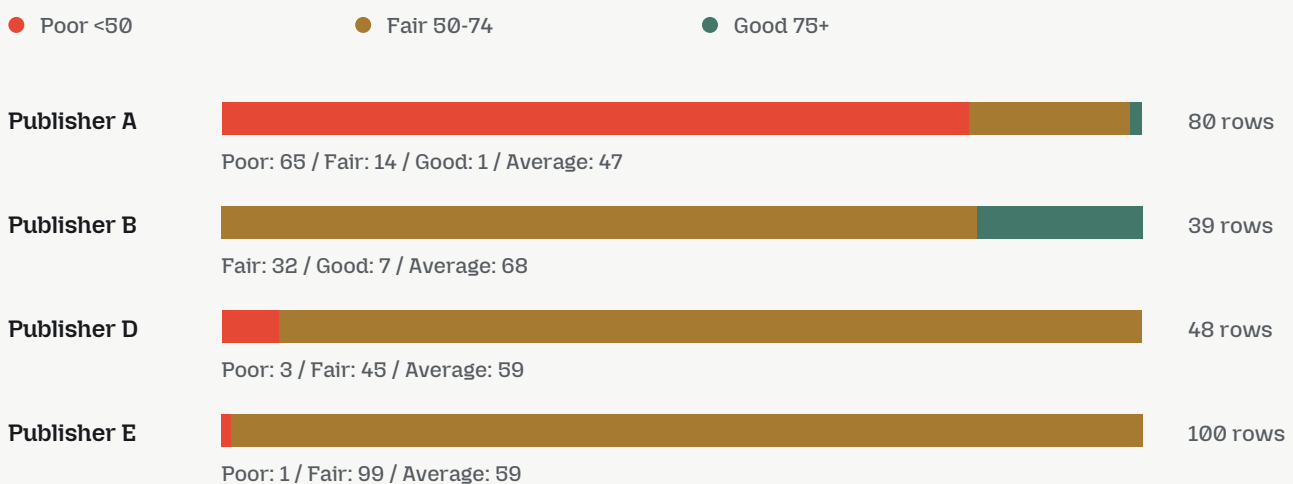
Every available page record is shown on the following maps. The two measures describe different things; they share a stable anonymous page ID.

Page Quality Map



Quality labels are source AI-model assessments against Search Quality Rater Guidelines. They are not Google ratings or independent editorial verification by Tityra.

Machine Readability Map



Machine Parse Score is the source's deterministic measure of served-HTML parsing. It is separate from Served-HTML Structure, Agent Readability and semantic-container scores. C provides no page-map records.

PAGE MAP / SOURCE A / 1 OF 3

Publisher A / page map

80 map rows / 81 pages in the audit scope. Quality order: Not rated, then lowest to highest.

PAGE ID	SOURCE QUALITY RATING						PARSE SCORE / 100			
	Not rated	Lowest	Low	Medium	High	Highest	0	50	75	100
A-069			●			Low				52
A-071			●			Low				53
A-074			●			Low				54
A-001				●		Medium				41
A-002				●		Medium				41
A-003				●		Medium				41
A-005				●		Medium				42
A-006				●		Medium				42
A-007				●		Medium				42
A-008				●		Medium				42
A-009				●		Medium				42
A-010				●		Medium				42
A-011				●		Medium				42
A-012				●		Medium				42
A-013				●		Medium				42
A-014				●		Medium				43
A-015				●		Medium				43
A-016				●		Medium				43
A-017				●		Medium				43
A-018				●		Medium				43
A-019				●		Medium				43
A-020				●		Medium				43
A-021				●		Medium				43
A-023				●		Medium				43
A-024				●		Medium				43
A-025				●		Medium				43
A-026				●		Medium				43

Source A pp63-71. IDs follow row order in the original page-rating register, with URLs withheld. Exact source page and row references are embedded in the evidence register.

The map denominator differs from the full audit count. The source does not explain all exclusions; no missing page result is invented here.

PAGE MAP / SOURCE A / 2 OF 3

Publisher A / page map

80 map rows / 81 pages in the audit scope. Quality order: Not rated, then lowest to highest.

PAGE ID	SOURCE QUALITY RATING						PARSE SCORE / 100			
	Not rated	Lowest	Low	Medium	High	Highest	0	50	75	100
A-027				●		Medium				43
A-029				●		Medium				43
A-030				●		Medium				43
A-031				●		Medium				43
A-032				●		Medium				44
A-033				●		Medium				44
A-034				●		Medium				44
A-035				●		Medium				44
A-036				●		Medium				44
A-037				●		Medium				45
A-038				●		Medium				45
A-039				●		Medium				45
A-040				●		Medium				45
A-042				●		Medium				45
A-045				●		Medium				46
A-046				●		Medium				46
A-047				●		Medium				47
A-048				●		Medium				47
A-049				●		Medium				47
A-050				●		Medium				47
A-051				●		Medium				47
A-052				●		Medium				47
A-053				●		Medium				47
A-054				●		Medium				47
A-055				●		Medium				47
A-056				●		Medium				47
A-057				●		Medium				48

Source A pp63-71. IDs follow row order in the original page-rating register, with URLs withheld. Exact source page and row references are embedded in the evidence register.

The map denominator differs from the full audit count. The source does not explain all exclusions; no missing page result is invented here.

PAGE MAP / SOURCE A / 3 OF 3

Publisher A / page map

80 map rows / 81 pages in the audit scope. Quality order: Not rated, then lowest to highest.

PAGE ID	SOURCE QUALITY RATING						PARSE SCORE / 100			
	Not rated	Lowest	Low	Medium	High	Highest	0	50	75	100
A-058				●			Medium			48
A-059				●			Medium			48
A-060				●			Medium			48
A-061				●			Medium			48
A-062				●			Medium			49
A-063				●			Medium			49
A-064				●			Medium			49
A-065				●			Medium			49
A-066				●			Medium			50
A-067				●			Medium			51
A-068				●			Medium			51
A-070				●			Medium			53
A-072				●			Medium			53
A-073				●			Medium			54
A-075				●			Medium			55
A-076				●			Medium			57
A-078				●			Medium			67
A-004					●		High			41
A-022					●		High			43
A-028					●		High			43
A-041					●		High			45
A-043					●		High			45
A-044					●		High			46
A-077					●		High			61
A-079					●		High			67
A-080					●		High			77

Source A pp63-71. IDs follow row order in the original page-rating register, with URLs withheld. Exact source page and row references are embedded in the evidence register.

The map denominator differs from the full audit count. The source does not explain all exclusions; no missing page result is invented here.

PAGE MAP / SOURCE B / 1 OF 2

Publisher B / page map

39 map rows / 40 pages in the audit scope. Quality order: Not rated, then lowest to highest.

PAGE ID	SOURCE QUALITY RATING						PARSE SCORE / 100			
	Not rated	Lowest	Low	Medium	High	Highest	0	50	75	100
B-002			●			Low				56
B-004			●			Low				58
B-005			●			Low				59
B-007			●			Low				60
B-008			●			Low				60
B-010			●			Low				61
B-011			●			Low				61
B-012			●			Low				61
B-013			●			Low				62
B-014			●			Low				62
B-016			●			Low				63
B-017			●			Low				70
B-018			●			Low				70
B-019			●			Low				71
B-020			●			Low				71
B-021			●			Low				72
B-022			●			Low				72
B-023			●			Low				72
B-024			●			Low				73
B-026			●			Low				74

Source B pp54-55. IDs follow row order in the original page-rating register, with URLs withheld. Exact source page and row references are embedded in the evidence register.

The map denominator differs from the full audit count. The source does not explain all exclusions; no missing page result is invented here.

PAGE MAP / SOURCE B / 2 OF 2

Publisher B / page map

39 map rows / 40 pages in the audit scope. Quality order: Not rated, then lowest to highest.

PAGE ID	SOURCE QUALITY RATING						PARSE SCORE / 100			
	Not rated	Lowest	Low	Medium	High	Highest	0	50	75	100
B-029			●				Low			74
B-030			●				Low			74
B-031			●				Low			74
B-032			●				Low			74
B-037			●				Low			76
B-038			●				Low			76
B-039			●				Low			76
B-006				●			Medium			59
B-009				●			Medium			60
B-015				●			Medium			62
B-025				●			Medium			73
B-027				●			Medium			74
B-028				●			Medium			74
B-033				●			Medium			75
B-034				●			Medium			75
B-035				●			Medium			75
B-001					●		High			54
B-003					●		High			58
B-036					●		High			75

Source B pp54-55. IDs follow row order in the original page-rating register, with URLs withheld. Exact source page and row references are embedded in the evidence register.

The map denominator differs from the full audit count. The source does not explain all exclusions; no missing page result is invented here.

PAGE MAP / SOURCE D / 1 OF 2

Publisher D / page map

48 map rows / 52 pages in the audit scope. Quality order: Not rated, then lowest to highest.

PAGE ID	SOURCE QUALITY RATING						PARSE SCORE / 100			
	Not rated	Lowest	Low	Medium	High	Highest	0	50	75	100
D-002	●						Not rated			41
D-001			●				Low			41
D-009			●				Low			56
D-017			●				Low			58
D-022			●				Low			60
D-035			●				Low			63
D-003				●			Medium			46
D-004				●			Medium			52
D-005				●			Medium			53
D-006				●			Medium			54
D-007				●			Medium			54
D-008				●			Medium			56
D-010				●			Medium			57
D-011				●			Medium			57
D-013				●			Medium			57
D-014				●			Medium			57
D-015				●			Medium			57
D-016				●			Medium			58
D-018				●			Medium			58
D-019				●			Medium			59
D-020				●			Medium			59
D-021				●			Medium			60
D-023				●			Medium			60
D-024				●			Medium			60

Source D pp55-57. IDs follow row order in the original page-rating register, with URLs withheld. Exact source page and row references are embedded in the evidence register.

The map denominator differs from the full audit count. The source does not explain all exclusions; no missing page result is invented here.

PAGE MAP / SOURCE D / 2 OF 2

Publisher D / page map

48 map rows / 52 pages in the audit scope. Quality order: Not rated, then lowest to highest.

PAGE ID	SOURCE QUALITY RATING						PARSE SCORE / 100			
	Not rated	Lowest	Low	Medium	High	Highest	0	50	75	100
D-025				●			Medium			61
D-026				●			Medium			61
D-029				●			Medium			62
D-030				●			Medium			62
D-031				●			Medium			62
D-032				●			Medium			62
D-033				●			Medium			62
D-034				●			Medium			63
D-036				●			Medium			63
D-037				●			Medium			63
D-038				●			Medium			63
D-039				●			Medium			63
D-043				●			Medium			65
D-045				●			Medium			65
D-046				●			Medium			66
D-048				●			Medium			67
D-012					●		High			57
D-027					●		High			62
D-028					●		High			62
D-040					●		High			64
D-041					●		High			64
D-042					●		High			64
D-044					●		High			65
D-047					●		High			66

Source D pp55-57. IDs follow row order in the original page-rating register, with URLs withheld. Exact source page and row references are embedded in the evidence register.

The map denominator differs from the full audit count. The source does not explain all exclusions; no missing page result is invented here.

PAGE MAP / SOURCE E / 1 OF 3

Publisher E / page map

100 map rows / 100 pages in the audit scope. Quality order: Not rated, then lowest to highest.

PAGE ID	SOURCE QUALITY RATING						PARSE SCORE / 100			
	Not rated	Lowest	Low	Medium	High	Highest	0	50	75	100
E-004	●						Not rated			51
E-003			●				Low			50
E-024			●				Low			56
E-041			●				Low			58
E-052			●				Low			59
E-068			●				Low			61
E-076			●				Low			62
E-083			●				Low			63
E-084			●				Low			63
E-086			●				Low			64
E-087			●				Low			64
E-089			●				Low			64
E-097			●				Low			67
E-007				●			Medium			52
E-008				●			Medium			52
E-009				●			Medium			52
E-011				●			Medium			53
E-018				●			Medium			55
E-019				●			Medium			55
E-020				●			Medium			55
E-022				●			Medium			56
E-026				●			Medium			56
E-037				●			Medium			57
E-040				●			Medium			58
E-048				●			Medium			58
E-058				●			Medium			59
E-066				●			Medium			61
E-067				●			Medium			61
E-075				●			Medium			61
E-077				●			Medium			62
E-088				●			Medium			64
E-090				●			Medium			64
E-091				●			Medium			65
E-093				●			Medium			66

Source E pp63-67. IDs follow row order in the original page-rating register, with URLs withheld. Exact source page and row references are embedded in the evidence register.

PAGE MAP / SOURCE E / 2 OF 3

Publisher E / page map

100 map rows / 100 pages in the audit scope. Quality order: Not rated, then lowest to highest.

PAGE ID	SOURCE QUALITY RATING						PARSE SCORE / 100			
	Not rated	Lowest	Low	Medium	High	Highest	0	50	75	100
E-094				●			Medium			66
E-096				●			Medium			66
E-098				●			Medium			67
E-099				●			Medium			67
E-100				●			Medium			68
E-001					●		High			49
E-002					●		High			50
E-005					●		High			51
E-006					●		High			52
E-010					●		High			52
E-012					●		High			53
E-014					●		High			54
E-015					●		High			54
E-016					●		High			54
E-017					●		High			54
E-021					●		High			56
E-025					●		High			56
E-027					●		High			56
E-028					●		High			56
E-029					●		High			56
E-030					●		High			56
E-031					●		High			57
E-032					●		High			57
E-033					●		High			57
E-034					●		High			57
E-035					●		High			57
E-036					●		High			57
E-038					●		High			58
E-042					●		High			58
E-043					●		High			58
E-044					●		High			58
E-045					●		High			58
E-046					●		High			58
E-047					●		High			58

Source E pp63-67. IDs follow row order in the original page-rating register, with URLs withheld. Exact source page and row references are embedded in the evidence register.

PAGE MAP / SOURCE E / 3 OF 3

Publisher E / page map

100 map rows / 100 pages in the audit scope. Quality order: Not rated, then lowest to highest.

PAGE ID	SOURCE QUALITY RATING						PARSE SCORE / 100			
	Not rated	Lowest	Low	Medium	High	Highest	0	50	75	100
E-049					●	High				58
E-050					●	High				58
E-051					●	High				58
E-053					●	High				59
E-054					●	High				59
E-055					●	High				59
E-056					●	High				59
E-057					●	High				59
E-059					●	High				59
E-060					●	High				59
E-061					●	High				59
E-062					●	High				60
E-063					●	High				60
E-064					●	High				60
E-069					●	High				61
E-070					●	High				61
E-071					●	High				61
E-072					●	High				61
E-073					●	High				61
E-074					●	High				61
E-078					●	High				62
E-079					●	High				62
E-080					●	High				62
E-081					●	High				62
E-082					●	High				62
E-092					●	High				65
E-095					●	High				66
E-013						● Highest				53
E-023						● Highest				56
E-039						● Highest				58
E-065						● Highest				60
E-085						● Highest				63

Source E pp63-67. IDs follow row order in the original page-rating register, with URLs withheld. Exact source page and row references are embedded in the evidence register.

07 / MX FINDINGS

Response failures need URL-level investigation

01 First: verify delivery

MX finding / Source evidence and proposed verification

Observation

A records 19 non-200 outcomes among 100 requested URLs; B records one among 41. D records zero among 52 and E zero among 100. These request inventories have different denominators from the pages successfully assessed.

Evidence

Source A pp35-36 and B p26 list the response outcomes. D p27 and E pp34-35 provide the comparison inventories. A's executive summary separately reports seven 5xx probe errors; that is a different probe count, not a substitute for the 19 request outcomes.

Potential impact

A failed response can prevent a consumer from obtaining the intended page during that request. A redirect may be correct canonicalisation; a non-200 total is not itself a count of broken content pages.

Recommendation

Classify the failed endpoints by purpose, final destination and body. Check repeated failures against server logs and access controls. Preserve correct redirects and appropriate error responses.

Verification

Re-request each affected URL at controlled intervals, retaining status, headers, redirect chain and response body. Compare with a valid-page and missing-page control. Establish whether an outcome persists before sizing remediation.

08 / MX FINDINGS

Served and rendered content need comparison

02 Verify the article body

MX finding / Source evidence and proposed verification

Observation

The source pipeline flags served-versus-rendered differences on 5/81 pages in A, 40/40 in B, 51/52 in D and 100/100 in E. Maximum reported gap scores are 95, 97, 96 and 92 respectively. C supplies no detailed gap register.

Evidence

A p51; B p42; D pp43-44; E p51. These are source "SPA Shell" flags. Paired response/browser snapshots and the referenced per-page CSVs were not supplied. The flags do not establish that every affected response is empty.

Potential impact

A consumer that does not render the page may receive less of the intended article or recipe. That impact depends on which content is missing and on the consumer's capabilities; this test does not describe the behaviour of every named AI product.

Recommendation

Check the main content in paired views of representative templates. Distinguish real omissions from navigation, consent overlays, delayed content and scanner extraction behaviour before deciding on server rendering or other changes.

Verification

Capture raw HTML and the browser view for the same URL and access conditions. Compare the actual article text, headings, author, date and relevant entities. Record exact missing content and re-run after changes.

09 / MX FINDINGS

Structured-data coverage varies by template

03 Review relevant entities

MX finding / Source evidence and proposed verification

Observation

The full reports record JSON-LD on 1/81 pages in A, 1/40 in B, 24/52 in D and 100/100 in E. D flags 88 property gaps; E flags 55 WebPage name gaps. C reports a score but no inventory.

Evidence

A pp45-47; B pp36-38; D pp36-38; E pp44-46. D groups four recommended WebSite properties on 22 pages each. E groups the name flag on 55 pages. Raw entities and per-entity CSVs are not supplied.

Potential impact

Sparse or inconsistent entity descriptions may leave a consumer to infer authors, publishers, page identity or relationships. Lack of JSON-LD does not mean there are no usable facts in HTML.

Recommendation

Model the entities actually present on the chosen templates and keep values consistent with visible content. Add relevant fields and stable identifiers for intended consumers. Schema.org is a vocabulary, not a universal list of mandatory properties.

Verification

Parse the actual JSON-LD, validate vocabulary and relevant consumer requirements, and compare values with the visible page. Check representative instances after template changes. Reconcile the score calculation separately.

10 / MX FINDINGS

Discovery inventory needs reconciliation

04 Confirm the crawl boundary

MX finding / Source evidence and proposed verification

Observation

A records no readable sitemap. B's summary lists 443 URLs while its expanded-index narrative describes about 3.9 million. D lists 30317 URLs alongside a much larger shard narrative. E lists 30421. C lists 36149 but only one page was reached.

Evidence

A p41; B pp31-32; D p32; E p40; C pp2-3. Different host, index or filtering scopes could explain some counts, but the underlying index files and crawl configuration were not supplied.

Potential impact

An unclear inventory prevents a defensible coverage percentage and can misdirect remediation. It does not by itself prove that a publisher's content is low quality or that a search engine discounts it.

Recommendation

Reconcile sitemap indexes, nested files, hosts, canonical URLs and the crawl filter. Identify which URLs the organisation intends to publish and which are intentionally excluded. Do not add noindex or non-content resources simply to satisfy a scanner count.

Verification

Save and parse the original index and child sitemaps; deduplicate by the agreed canonical rule. Match the scored-page and requested-URL lists to that inventory. Publish coverage only when numerator and denominator share the same scope.

11 / MX FINDINGS

Page regions need explicit structure

05 Inspect representative layouts

MX finding / Source evidence and proposed verification

Observation

The detailed semantic-container scores are A 15 served / 15 rendered, B 39/50, D 26/38 and E 60/40. They describe a different rubric from the headline served-HTML scores and include worst-page examples.

Evidence

A p52; B p44; D p45; E p52. The source records deep generic-container chains, but those statistics alone do not establish missing article content or identify the framework that caused them.

Potential impact

Poor landmarks and heading relationships can make it harder to separate main content from navigation or repeated modules. A bare div is not automatically a defect, and a changed container ratio does not prove better comprehension.

Recommendation

Inspect the document outline and accessibility tree on article, listing, author and recipe templates. Prefer appropriate native main, article, nav, header and footer elements. Use ARIA where needed, avoiding redundant roles.

Verification

Confirm that primary content and page regions are identifiable in served HTML and the browser tree. Check meaningful headings and reading order. Verify the same repeated pattern on other instances before declaring a site-wide template issue.

12 / MX FINDINGS

Simulated crawler responses need policy context

06 Keep permission and access separate

MX finding / Source evidence and proposed verification

Observation

All eight request profiles in A, B, D and E received HTTP 200 on the homepage. Each test used supplied User-Agent strings; it did not authenticate the actual crawler.

Evidence

A p33; B p24; D p26; E p32. C supplies no simulated-crawler table. A's robots excerpt includes a GPTBot-specific disallow rule, illustrating why an HTTP success response cannot establish intended crawl permission.

Potential impact

A simulated request may receive a page while the real crawler is restricted by policy, IP-based rules or other checks. A success status also says nothing about whether the body contains the intended article.

Recommendation

Rename the check "Simulated AI Crawler Access Tests". Compare bot-specific rules with the publisher's intended policy. Do not remove deliberate exclusions just to make an access table look successful.

Verification

Review complete robots rules, response bodies and verified crawler traffic where available. Retain the test environment and headers. Distinguish an allowed simulation, a permitted crawl and a verified real-bot visit.

13 / MX FINDINGS

One reached page cannot represent a site

07 Expand scope before conclusions

MX finding / Source evidence and proposed verification

Observation

C is a four-page thin-site summary. It records one page examined out of a sitemap listing 36149 URLs. Its headline scores therefore describe that reached page, not the wider publisher.

Evidence

Source C pp2-4 names a possible path-filter mismatch, crawl ceiling or locale restriction. Those are possible explanations; the supplied summary does not establish which one caused the restricted crawl.

Potential impact

Using the single-page result as a site-wide verdict would overstate coverage and could prioritise changes on unrepresented templates. It could also confuse a crawl configuration issue with a content-access defect.

Recommendation

Agree the intended host and path scope, review filters and exclusions, then select representative pages. Keep C's reported results as a single-page observation until a broader sample has been collected.

Verification

Retain the crawl configuration, discovery queue and fetch outcomes. Confirm coverage of agreed templates and explain exclusions. Only then update the scope statement and any wider conclusions.

14 / MX FINDINGS

Size warnings are heuristic signals

08 Measure actual content loss

MX finding / Source evidence and proposed verification

Observation

Truncation warnings are recorded on 81/81 pages in A, 40/40 in B, 47/52 in D and 3/100 in E. Their largest reported pages are 3206 KB, 261 KB, 624 KB and 318 KB. The scanner uses a 250 KB threshold and content-offset windows.

Evidence

A p51; B p42; D p43; E p51. These are configured scanner checks. The source also flags inline blocks on all sampled pages in each full report; that does not independently demonstrate failed article extraction.

Potential impact

A large response may cost more to transfer or parse and could exceed a particular consumer's limit. The supplied reports do not demonstrate truncation by an actual consumer or a universal 250 KB limit.

Recommendation

Measure transferred size and the position of important content. Check whether repeated modules or large inline blocks obstruct the target consumer. Benchmark any change instead of blindly externalising every inline block.

Verification

Use the intended consumer or an agreed bounded-fetch test to capture the received bytes and extracted main content. Confirm whether any text is lost, then repeat under the same conditions after optimisation.

15 / EVIDENCE REGISTER

Delivery evidence register

Recorded check	A	B	C	D	E
Requested URLs	100	41	Not stated	52	100
Non-200 outcomes	19	1	Not stated	0	0
Redirected addresses	4	3	Not stated	12	66
Served/rendered flags	5/81	40/40	Not stated	51/52	100/100
Truncation flags	81/81	40/40	Not stated	47/52	3/100
Missing-page control	404	404	Not supplied	404	404

Request inventory: A pp35-36, B p26, D p27, E pp34-35. Pipeline: A p51, B p42, D pp43-44, E p51. Controls: A p37, B p27, D p28, E p36.

Positive result: correct missing-page signals

The four full reports record HTTP 404 for their deliberately missing control URL, with no content-schema claim on that error page. A control result is evidence for that test; it is not proof that every missing path on each site behaves correctly.

Redirects need interpretation

E's 66 redirects include canonical slash variants in the source list. Redirects are not all defects. A's pipeline cross-host register reports 1/81 while its explanatory text names four; this discrepancy is recorded rather than silently choosing a count.

16 / EVIDENCE REGISTER

Entity and template evidence

Source	JSON-LD coverage	Types recorded	Property flags
A	1 of 81	2	No property gaps reported within the scanner's rules
B	1 of 40	4	No property gaps reported within the scanner's rules
C	Not supplied	Not supplied	Not supplied
D	24 of 52	15	88: four recommended WebSite fields on 22 pages each
E	100 of 100	14	55: WebPage name on 55 pages

Sources: A pp45-47, B pp36-38, D pp36-38, E pp44-46. C supplies only a summary score. A source claim of no checked property gaps does not establish complete entities when most sampled pages carry no JSON-LD.

Use the right entity for the content

Article and publisher context, an author identity and a recipe's properties serve different purposes. Check which entities the page actually describes. Add fields that match visible, attributable information; do not populate an article's date or author on an unrelated entity solely because a scanner recommends it.

What the flags mean

The source uses required and recommended bands from its own consumer-oriented rubric. Schema.org itself does not make every flagged field mandatory. Rich-result eligibility needs the relevant vendor test; it cannot be inferred from an SDQ total.

17 / EVIDENCE REGISTER

Source reconciliation register

Source issue	Recorded values	Treatment here
SDQ component display	A: sum 0 / total 1; B: sum 1 / total 2; D: sum 40 / total 38; E: sum 73 / total 73	Keep source aggregates; ask for unrounded values and aggregation method
Pipeline summary	A and E say no survivability issues despite failed checks	Do not repeat the issue-free conclusion
Cross-host count in A	Pipeline label 1/81; explanatory text four pages	No corrected count invented
Page-map coverage	A: 80 map rows / 81 assessed; B: 39/40; D: 48/52	Retain the map rows; exclusion rationale remains unverified
Sitemap scope in B/D	Summary counts differ from expanded-index narrative	Do not calculate estate coverage or endorse the bulk-content judgement
Content uniqueness in D	Two distinctive / 50 repeated; later text says remaining 42 are distinctive or expected boilerplate	Do not publish a duplication verdict from the inconsistent narrative
Response timing in E	Initial 1888 ms; source healthy threshold 1500 ms; narrative says healthy in both views	Initial result described as elevated under the source threshold

Sources: A p45 and p51; B p36 and pp31-32; D p37, p32 and p52; E pp44-45, p51 and p39. Map counts: A/B/D p3 and p5. Display sums are arithmetic of printed components, not recalculated audit scores.

These checks review the supplied reports' internal consistency. They are not a new scan and do not independently validate the scanner. Unresolved evidence stays visible rather than being replaced with plausible-looking data.

18 / SUPPORTING TECHNICAL FINDINGS

Supporting Technical Findings

Accessibility, performance and SEO provide useful context. They do not displace the content-access and structure questions at the centre of this MX specimen.

Recorded score / 100	A	B	C*	D	E
Automated accessibility	55	61	65	53	63
Performance	85	85	55	70	85
SEO	76	87	91	81	93

Sources: A/B/D/E pp3-4; C pp3-4. *C is one page. These are source scores; no peer-ranking dataset is supplied and no new comparison grade is assigned.

Accessibility needs contextual checks

Automated tools can flag contrast, labels, image alternatives and document structure. The report scores do not certify WCAG conformance or establish legal non-compliance. Validate affected components and include keyboard, reading-order and assistive-technology checks where relevant.

SEO is not a visibility guarantee

A high SEO score does not demonstrate indexing, search rankings or citations by AI systems. Keep technical observations distinct from actual audience or crawler outcomes.

19 / SUPPORTING TECHNICAL FINDINGS

Security and response policies

Recorded transport / header	A	B	C	D	E
HTTPS transport	Yes	Yes	Not detailed	Yes	Yes
HSTS	Yes	No	Not detailed	No	Yes
Content-Security-Policy	Yes	No	Not detailed	No	No
X-Frame-Options	Yes	No	Not detailed	No	No
X-Content-Type-Options	Yes	No	Not detailed	No	No

Sources: A p53, B p45, D p46, E p53. C p4 reports three of five types but does not identify them; no individual header values are inferred. HTTPS is transport, not an additional HTTP header.

Presence is only the starting point

The source records header presence, not the correctness or effectiveness of every value. A policy may need route-specific exceptions, application testing or a phased rollout. Review header values and browser behaviour before describing the site as protected.

Recommended verification

Capture headers on representative valid pages, redirects and error responses. Validate each policy against the application and test for regressions. A Content-Security-Policy change is not automatically a safe one-line fix.

20 / SUPPORTING TECHNICAL FINDINGS

Response timing observations

Source	Initial slowest request	Three repeat requests	Repeat median
A	4925 ms	2518 / 228 / 710 ms	710 ms
B	8543 ms	722 / 1633 / 447 ms	722 ms
D	8311 ms	411 / 3314 / 746 ms	746 ms
E	1888 ms	1277 / 657 / 693 ms	693 ms

Sources: A p40, B p30, D p31, E p39. C does not provide repeat-request timing evidence. These are response observations, distinct from the source's Machine Processing Speed metric.

What the timing suggests

The initial and repeated requests vary substantially in several sources. A, B and D exceed the source's 3000 ms slow threshold initially. E's 1888 ms falls in its acceptable-but-elevated range, while its repeat median is below 1500 ms.

What it does not prove

The reports call the initial request cold-cache and repeat requests returning-visitor views. No cache-state record is supplied, so this specimen does not attribute the variation to a cold origin or claim these values reproduce every visitor's experience.

Re-test from a recorded location with controlled cache and request conditions. Preserve each sample rather than replacing variable requests with a single reassuring median.

Optional conventions and experiments

llms.txt and alternative formats

The llms.txt project is a proposal for agent-oriented guidance. It may be useful to chosen consumers, but its absence is not a mandatory failure. B's discovery section both describes an llms.txt file and says it was not found; its presence is therefore unresolved in the supplied evidence. Do not implement the source's HTML-wrapper recommendation without checking the intended format and consumer.

MX markers and maturity models

The source scoring and maturity models include MX-specific metadata associated with The Gathering. These optional fields are not owned by Tityra and are not prerequisites for accessible HTML. No new score penalty is assigned for their absence in this edition.

Agent cards, content-origin markers and Markdown

An agent-service card is relevant when an organisation offers that service. Content-origin declarations need an actual editorial or provenance use case. Optional Markdown negotiation must preserve important information for the consuming system. None of these techniques proves successful crawling, citations or trustworthy content on its own.

Source context: B pp32-33 and the source methodology sections. Proposal reference: <https://llmstxt.org/>. The page's status as a proposal was checked for this edition; that is separate from the October source audits.

22 / RE-TEST PLAN

Re-test method and acceptance criteria

Test area	Evidence to retain	Acceptance decision
Article delivery	Raw response, rendered snapshot, exact content comparison	Intended body and metadata reach the agreed consumer
Response handling	Status, final URL, headers and body; missing-path control	Correct distinction between valid, restricted, moved and absent resources
Structured data	Parsed entities, visible values and consumer validation	Relevant fields correct; entities and identities coherent
Discovery / policy	Full robots file, sitemap tree, canonical inventory	Intended scope and bot permissions explicit
Report quality	URL denominator, component calculations, source references	Counts reconcile; limitations and outstanding evidence stated

These are proposed verification steps, not tests performed for this edition. A re-test should record the date, environment, credentials or consent state where relevant, request headers and the exact target sample.

HTML and PDF acceptance remain separate. If a PDF is discovered and reviewed, identify that document and the checks actually performed. Additional documents require a separate agreed review. Tags or an embedded conformance identifier alone are not PDF/UA certification.

23 / SOURCES

Source traceability and attribution

This specimen replaces publisher identities with A-E and omits real domains, paths and article titles. It selects recorded evidence from five independent reports; it is not a single client audit and includes no fabricated scan results.

Source	Original pages	SHA-256 of supplied PDF
Publisher A	84	12daf7a3bbbf621b2212c7543fc1117b6ffad23ae6a35cbd0cc16c3bc7dbe452
Publisher B	68	ff776bcd62d4af8661489b4539387bea2bddd5179029479ade69fa9c8a5e98a7
Publisher C	4	352f171e281de921ac9bcf5fc1d57212d4573de65390adcf4f36364fa700f14a
Publisher D	70	bcc0fdf5a6839cb54b707a03f07252f9c383e6fe8ee4f400a944973c62f1ad82
Publisher E	80	4a00eca373890f5523ae7a3906f577d4421cc1fcf669f84e251d9e769aa2544e

Roles and dates

The source covers name Tom Cranstoun; the four full reports identify the Web Audit Suite, CogNovaMX and an audit date of 7 October 2026. C does not state a dated audit record. Tityra is the presentation and service brand of Digital Domain Technologies Ltd, trading as CogNovaMX; no separate unverified partner role is assigned.

The anonymous evidence register is embedded as specimen-evidence.json. Source page references refer to the original PDFs, not this edition's page numbers. Original reports and private provenance payloads are not embedded in the public specimen.

24 / SOURCES

Technical references

HTTP and crawler policy

Google crawler identification and request verification. A User-Agent string alone does not authenticate the crawler.

<https://developers.google.com/crawling/docs/crawlers-fetchers/overview-google-crawlers>

Google-Extended

This is a robots product token, not a distinct HTTP crawler identity. The simulated request profile is not a real Google-Extended visit.

<https://developers.google.com/crawling/docs/crawlers-fetchers/google-common-crawlers>

Rendered content

Google Search documents crawling and JavaScript rendering as separate phases; consumers differ in rendering capability.

<https://developers.google.com/search/docs/crawling-indexing/javascript/javascript-seo-basics>

Sitemap fields

changefreq and priority are optional. Their absence is not, by itself, a machine-readability failure.

<https://www.sitemaps.org/protocol.html>

lms.txt

An optional proposal for guidance to agents, considered here as an experiment rather than a universal requirement.

<https://lms.txt.org/>

Schema.org

Vocabulary reference. Check the intended consumer's requirements separately.

<https://schema.org/docs/schemas.html>

WCAG

The source scores use automated checks; specialist evaluation is needed for conformance claims.

<https://www.w3.org/TR/WCAG21/>

YOUR NEXT AUDIT

Start with the content that matters

Agree a representative sample of pages, templates and documents. An MX Audit examines the content software receives and gives your team documented findings and practical recommendations.



Discuss an MX Audit

info@tityra.com

tityra.com

Digital Domain Technologies Ltd, trading as CogNovaMX.