Taggart remediation report — sample

Generated by taggart/1.0.0 · profile pdfua1 · 2026-10-02T00:08:12.607934+00:00

Conformance statement

This file does not conform to the selected profile: 1 machine-checkable failure(s) remain.

7 findings decided — 7 repaired. 1 remains in the output, recorded below with the decision made about it.

Conditions Taggart does not check

39 conditions in this profile are not decided by Taggart: 27 need a person's judgement and no software can decide them; 12 could be checked by software and Taggart does not check them yet. They produce no finding, so the counts below say nothing about them. 39 of them have not been confirmed by a person.

01-006 Structure type is not semantically appropriate (not confirmed); 02-002 Role mapping is semantically inappropriate (not confirmed); 03-001 One or more Actions lead to flickering (not confirmed); 03-002 Multimedia objects contain flickering content (not confirmed); 03-003 JavaScript actions lead to flickering (not confirmed); 04-001 Information conveyed by contrast, colour, format or layout is not tagged (not confirmed); 05-001 Media annotation present, but audio content not available in another form (not confirmed); 05-002 Audio annotation present, but content not available in another form (not confirmed); 05-003 JavaScript uses beep but provides no other notification (not confirmed); 08-001 OCR-generated text contains significant errors (not confirmed); 08-002 OCR-generated text is not tagged (not confirmed); 09-002 Structure elements are nested in a semantically inappropriate manner (not confirmed); 09-003 Structure type is not semantically appropriate after role mapping (not confirmed); 12-001 Stretched characters are not represented appropriately (not confirmed); 13-002 A link with a meaningful background lacks alternative text for both (not confirmed); 13-005 ActualText used for a Figure for which alternative text is more appropriate (not confirmed); 13-006 A graphic with meaning only as part of a group is tagged on its own (not confirmed); 13-007 A more accessible representation is not used (not confirmed); 13-008 ActualText not present when a Figure is intended to be consumed primarily as text (not confirmed); 14-004 Numbered heading tags do not use Arabic numerals (not confirmed); 14-005 Content representing a 7th level (or higher) heading does not use an H7 (or higher) tag (not confirmed); 15-005 A given cell's header cannot be unambiguously determined (not confirmed); 22-001 Article threads do not reflect logical reading order (not confirmed); 28-003 An annotation is used for visual formatting but is not tagged according to its semantic function (not confirmed); 28-013 An IsMap entry is true but the functionality is not provided in some other way (not confirmed); 29-001 A script requires specific timing for individual keystrokes (not confirmed); 31-010 A font program is embedded that is not legally embeddable for unlimited, universal rendering (not confirmed); 31-011 The embedded font program does not contain glyphs for all referenced glyphs (not confirmed); 31-012 A glyph present in an embedded Type 1 font program is not listed in CharSet (not confirmed); 31-013 A glyph listed in CharSet is not present in the Type 1 font program (not confirmed); 31-014 A glyph present in an embedded CID font program is not listed in CIDSet (not confirmed); 31-015 A glyph listed in CIDSet is not present in the CID font program (not confirmed); 31-016 Glyph widths in the font dictionary and the font program differ (not confirmed); 31-017 A non-symbolic TrueType font has no non-symbolic cmap (not confirmed); 31-018 A glyph cannot be looked up by any non-symbolic cmap entry (not confirmed); 31-023 Differences present but the font has no (3,1) Microsoft Unicode cmap (not confirmed); 31-025 The embedded font program for a symbolic TrueType font contains no cmap (not confirmed); 31-026 A symbolic TrueType font has multiple cmaps but none is a (3,0) Microsoft Symbol cmap (not confirmed); 31-030 Characters used in text showing operators reference the .notdef glyph (not confirmed)

Taggart reports what its check engine can verify. It does not and cannot determine legal compliance; that judgement belongs to a qualified human reviewer.

Files

Source	2-word-export-tagged.pdf
Source SHA-256	eb70aa4207178bca265ad1af3f57143e6bcda8fbd83b8c83437dcd337005df91
Pages	2
Output	sample-remediated.pdf
Output SHA-256	27d6b2ccec1102802c2636347e79519d92ef74713af7ca0bf8ca586705f6ab64

Findings

Severity	Count	Meaning
Blocker	1	Fails a machine-checkable PDF/UA-1 requirement. The file cannot claim PDF/UA-1 conformance until this is resolved.
Warning	0	Passes the machine check or has no machine test, but Taggart's heuristics indicate it very likely fails human review.
Needs a human decision	0	Requires a human decision. Taggart presents the evidence and records the decision; it never auto-passes these.

Fixes applied

Repair heading levels — applied

• page 1 StructTreeRoot[0]/Document[0]/H2[2]: <H2> -> <H1>

• page 1 StructTreeRoot[0]/Document[0]/H4[4]: <H4> -> <H2>

• page 1 StructTreeRoot[0]/Document[0]/H2[6]: <H2> -> <H1>

• page 1 StructTreeRoot[0]/Document[0]/H2[8]: <H2> -> <H1>

• page 1 StructTreeRoot[0]/Document[0]/H2[11]: <H2> -> <H1>

• page 2 StructTreeRoot[0]/Document[0]/H2[14]: <H2> -> <H1>

• page 2 StructTreeRoot[0]/Document[0]/H2[16]: <H2> -> <H1>

Repair right-to-left text to logical order — applied

• page 1 StructTreeRoot[0]/Document[0]/P[9]/Span[0]: set /ActualText to the logical-order text (94 chars)

• page 1 StructTreeRoot[0]/Document[0]/P[9]/Span[4]: set /ActualText to the logical-order text (19 chars)

• page 1 StructTreeRoot[0]/Document[0]/P[10]/Span[0]: set /ActualText to the logical-order text (76 chars)

• page 1 StructTreeRoot[0]/Document[0]/P[9]/Span[1]: set /Lang to he-IL

• page 1 StructTreeRoot[0]/Document[0]/P[9]/Span[3]: set /Lang to he-IL

• no content stream was modified: the rendered page is unchanged and the logical-order text is carried in /ActualText

Add alternative text to figures — applied

• page 2 StructTreeRoot[0]/Document[0]/P[13]/Figure[0]: set /Alt to 'Bar chart: defects by category. Untagged content 141 documen'

Build link structure elements — applied

• page 2: set link /Contents to 'ISO 14289-1 specification' (from the link's own text — review the wording)

Add annotation descriptions — applied

• page 2: set /Link /Contents

Set table headers, Scope and Headers references — applied

• page 1 StructTreeRoot[0]/Document[0]/Table[7]/TR[1]/TH[0]: <TD> -> <TH> (approved by the reviewer)

• page 1 StructTreeRoot[0]/Document[0]/Table[7]/TR[2]/TH[0]: <TD> -> <TH> (approved by the reviewer)

• page 1 StructTreeRoot[0]/Document[0]/Table[7]/TR[3]/TH[0]: <TD> -> <TH> (approved by the reviewer)

• page 1 StructTreeRoot[0]/Document[0]/Table[7]/TR[4]/TH[0]: <TD> -> <TH> (approved by the reviewer)

• page 1 StructTreeRoot[0]/Document[0]/Table[7]/TR[5]/TH[0]: <TD> -> <TH> (approved by the reviewer)

• page 1 StructTreeRoot[0]/Document[0]/Table[7]/TR[1]/TH[0]: set /Scope /Row

• page 1 StructTreeRoot[0]/Document[0]/Table[7]/TR[2]/TH[0]: set /Scope /Row

• page 1 StructTreeRoot[0]/Document[0]/Table[7]/TR[3]/TH[0]: set /Scope /Row

• page 1 StructTreeRoot[0]/Document[0]/Table[7]/TR[4]/TH[0]: set /Scope /Row

• page 1 StructTreeRoot[0]/Document[0]/Table[7]/TR[5]/TH[0]: set /Scope /Row

Create or repair the XMP metadata stream — not applied

Reason: conformance was requested but cannot be claimed, so the PDF/UA identifier was not written: re-validation found 1 blocker(s)

Human decisions

Rule	Decision	By	When	Justification
14-002	fixed	Sample reviewer	2026-10-02 00:08:12	The document's own numbering runs 1, 2, 3; the tags did not follow it.
14-003	fixed	Sample reviewer	2026-10-02 00:08:12	The document's own numbering runs 1, 2, 3; the tags did not follow it.
TG-RTL-001	fixed	Sample reviewer	2026-10-02 00:08:12	The Hebrew and Arabic passages were checked against the printed page and are stored pre-reversed.
13-004	fixed	Sample reviewer	2026-10-02 00:08:12	Description written from the chart's own data table, which appears above it on the same page.
28-012	fixed	Sample reviewer	2026-10-02 00:08:12	The link text already names the standard; the URI is an accurate description of where it goes.
28-004	fixed	Sample reviewer	2026-10-02 00:08:12	Written by hand: the annotation had no /Contents at all.
15-001	fixed	Sample reviewer	2026-10-02 00:08:12	The first cell of each row names the defect category that row is about, so it heads the row.

Findings in detail

06-002 — XMP metadata does not include the PDF/UA identifier (Blocker, 1 occurrence)

Matterhorn 06-002: The XMP metadata stream in the Catalog dictionary does not include the PDF/UA identifier.

The PDF/UA identification schema (pdfuaid:part) is absent, so the file does not actually claim PDF/UA-1 conformance even if it meets every other requirement. Taggart does not write the identifier as a batch fix: it is a claim of conformance, and it is written only through the conformance-claim step, which refuses unless re-validation of the output finds no blockers.

What a screen reader does: No direct effect. This is the machine-readable conformance claim that validators and procurement checks look for.

Maps to: PDF/UA-1 UA1:5

• The XMP metadata stream does not include the PDF/UA identification schema (pdfuaid:part), so the file does not claim PDF/UA conformance.

Method and provenance

Checks are expressed against the Matterhorn Protocol 1.1 (PDF Association): 31 checkpoints and 136 failure conditions, of which 87 can be decided by software and 47 require human judgement.

119 check(s) ran successfully.