Auswertung von 10.648 deutschen Websites: 94,1 Prozent nennen ihre Anschrift im Text - doch bei 73,0 Prozent fehlt maschinenlesbares Unternehmens-Markup

Rheine. Fast jedes untersuchte Unternehmen in Deutschland nennt seine Postanschrift ordentlich im sichtbaren Text oder im Impressum (94,1 %). Im technischen Quellcode fehlt jedoch auf knapp drei Vierteln der untersuchten Startseiten (73,0 % bzw. 7.778 von 10.648 Domains) ein strukturiertes Unternehmens-Markup nach Schema.org (Organization oder LocalBusiness). Selbst unter den 2.870 deutschen Websites, die ein solches Unternehmens-Markup im JSON-LD-Format einsetzen, hinterlegen nur 21,3 % (611 Domains) darin eine verschachtelte Postadresse (PostalAddress).

Das zeigt eine technische Untersuchung des Rheiner Software-Anbieters MIRDAT, für die im September und Oktober 2026 insgesamt 13.710 technisch erreichbare Websites aus öffentlichen Verzeichnissen (davon n = 10.648 in Deutschland) automatisiert analysiert wurden.

Wenig aktive Bot-Sperren - aber große Lücken in der Datenstruktur
In der Diskussion über KI-gestützte Suchsysteme wie ChatGPT Search, Perplexity oder Google Gemini wird häufig vermutet, dass Unternehmen automatisierte Abrufprogramme aktiv aussperren. Die Untersuchung der jeweiligen /robots.txt-Dateien bestätigt dies für deutsche Websites kaum: Lediglich 2,2 % der untersuchten deutschen Domains (239 von 10.648) blockieren gezielt KI-Such- und Retrieval-Crawler wie OAI-SearchBot, PerplexityBot, Applebot oder Claude-User (97,8 % bzw. 10.409 Domains ungesperrt). Etwas häufiger - nämlich bei 4,3 % (453 Domains) - werden reine Trainings-Crawler wie GPTBot, CCBot oder Google-Extended vom Einsammeln von Trainingsdaten ausgeschlossen. Eine pauschale Komplettsperre sämtlicher Suchmaschinen (User-agent: * mit Disallow: /) fand sich bei 0,6 % (60 Domains).

Technische Einordnung: Was fehlendes Markup bedeutet - und was nicht
MIRDAT weist ausdrücklich darauf hin, dass ein fehlendes Schema.org-Markup oder eine fehlende /llms.txt-Datei nicht automatisch bedeutet, dass ein Unternehmen in Suchmaschinen oder KI-Antworten unsichtbar ist. So dokumentiert etwa Google für seine KI-Suchfunktionen, dass weder spezielle KI-Textdateien noch gesonderte Schema-Auszeichnungen zwingende Voraussetzung für eine Aufnahme sind. Strukturierte JSON-LD-Auszeichnungen erleichtern automatisierten Systemen jedoch die eindeutige Zuordnung von Unternehmensname, Standort, Branche und Öffnungszeiten; bei 73,0 % der Domains liegen diese Angaben jedoch nicht zusätzlich in der geprüften strukturierten Form vor.

Bei dem noch jungen Dateiformat /llms.txt - einer kompakten Markdown-Übersicht für KI-Systeme - zeigte sich zudem ein deutlicher Unterschied zwischen oberflächlicher Erreichbarkeit und echtem Inhalt: Nachdem sogenannte Soft-404-Seiten herausgefiltert wurden, verfügten 4,7 % der untersuchten deutschen Websites (497 von 10.648) über eine valide Markdown-Datei (vor Filterung oberflächlich 9,7 % bzw. 1.030 Domains).

Internationaler Ländervergleich: Vorsprung englischsprachiger Märkte
Ein Blick über die Landesgrenzen verdeutlicht das Optimierungspotenzial: In Großbritannien (54,1 %, n = 839) und den USA (53,8 %, n = 2.114) setzt bereits mehr als jedes zweite Unternehmen auf maschinenlesbares Schema.org-Markup. Auch bei den maschinenlesbaren Postadressen liegen US-Websites mit 29,7 % (627 Domains) deutlich vor den deutschen Betrieben (6,5 % der Gesamtkohorte bzw. 21,3 % der Teilgruppe mit Markup).

"In den meisten Betrieben sind die Angaben zu Standort und Leistungen längst vorhanden - sie stehen im Impressum oder auf der Kontaktseite für menschliche Besucher. Im Quelltext fehlen jedoch häufig strukturierte Schema.org-Auszeichnungen wie die PostalAddress oder maschinenlesbare Übersichten wie die llms.txt. Wer seine Website technisch sauber strukturiert, nimmt automatisierten Such- und Antwortsystemen unnötige Interpretationsarbeit ab."

- Michael Röttger, Inhaber und Software-Entwickler von MIRDAT in Rheine

Zentrale Kennzahlen im Überblick (Deutschland, n = 10.648):

Anschrift im sichtbaren Text / Impressum:	**94,1 %**	Postadresse für Besucher erkennbar
Kein Schema.org-Unternehmens-Markup:	**73,0 %**	7.778 von 10.648 Startseiten ohne Organization/LocalBusiness
Schema.org-Unternehmens-Markup vorhanden:	**27,0 %**	2.870 von 10.648 Startseiten
Verschachtelte PostalAddress (in Entität):	**21,3 %**	611 von 2.870 Domains mit Markup
PostalAddress im JSON-LD (Gesamtkohorte):	**6,5 %**	695 von 10.648 erreichbaren DE-Websites
Valide /llms.txt vorhanden:	**4,7 %**	497 von 10.648 Domains (Soft-404-bereinigt)
KI-Suchcrawler in robots.txt gesperrt:	**2,2 %**	239 von 10.648 (97,8 % bzw. 10.409 ungesperrt)
KI-Trainings-Bots in robots.txt gesperrt:	**4,3 %**	453 von 10.648 Domains
Pauschalsperre aller Crawler (Disallow: /):	**0,6 %**	60 von 10.648 Domains

Über MIRDAT und den Entwickler:

Die Software MIRDAT wurde von Michael Röttger in Rheine (Münsterland) entwickelt. Sie entstand aus mehrjährigen Praxiserfahrungen an der Schnittstelle von gewerblichem Alltag, kaufmännischen Abläufen und Softwareentwicklung. MIRDAT führt automatisierte technische Außenprüfungen für Websites von kleinen und mittelständischen Unternehmen (KMU) durch und überwacht Auffindbarkeit, Strukturdaten und Systemintegrität.

Pressekontakt für Redaktionen & Rückfragen:

MIRDAT - Michael Röttger | Friedrich-Ebert-Ring 154 | 48429 Rheine (Münsterland / Kreis Steinfurt)
E-Mail: info@mirdat.de | Quellenseite, Grafiken & Datenblätter: https://www.mirdat.de/presse
Abdruck und Zitate honorarfrei bei Nennung von "Quelle: MIRDAT (mirdat.de/presse)"

Hinweise für Redaktionen & Bildmaterial:
- Grafiken & Rohdaten: Hochauflösende Vektorgrafiken (SVG/PNG) sowie tabellarische Factsheets stehen unter https://www.mirdat.de/presse zum Download bereit (Lizenz: CC BY 4.0).
- Zitierweise: 'Quelle: MIRDAT Website-Studie 2026 (mirdat.de/presse)'
- Regionalauswertungen: Aufbereitete Vergleichsdaten für Münsterland, Osnabrück, Euregio und deutsche Metropolregionen sind über das Presseportal frei abrufbar.