

# Wie die Erkennung funktioniert

Dieses Kapitel erklärt, wie Media Organizer zu seinem Urteil kommt. Sie brauchen das nicht für den Alltag — aber wenn Sie eine Einstufung nachvollziehen oder die Abdeckung Ihrer Website beurteilen wollen, ist es genau das Richtige.

## Zwei getrennte Schritte

Ein Scan besteht aus zwei Teilen, die bewusst getrennt sind:

1. **Inventar aufbauen.** Welche Mediendateien existieren überhaupt? Gelesen wird aus den konfigurierten Medien-Adaptern.
2. **Referenzen sammeln.** Wer verweist auf welche Datei? Das erledigen mehrere unabhängige *Suchbereiche*.

Erst danach vergleicht Media Organizer beides. Eine Datei ohne passende Referenz wird zum Kandidaten — mehr nicht.

Wichtig für das Verständnis: Ein Suchbereich **meldet nur Fundstellen**. Er urteilt nicht. Die Aussage „ohne Referenz“ trifft ausschließlich die auswertende Stelle, nachdem *alle* Suchbereiche gelaufen sind.

## Die Suchbereiche

| Bereich                | Was er durchsucht                                                                          |
|------------------------|--------------------------------------------------------------------------------------------|
| Artikel und Kategorien | Beiträge (Einleitungs- und Volltext, Beitragsbild) und Kategorien, inklusive Zusatzfelder. |

| Bereich          | Was er durchsucht                                                                                                                                             |
|------------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Module           | Das HTML von „Eigenes HTML“-Modulen, die Modul-Einstellungen und Ordner-Angaben. Auch unveröffentlichte Module und solche im Papierkorb zählen (einstellbar). |
| Versionshistorie | Ältere Fassungen von Beiträgen. Wer eine alte Version wiederherstellt, braucht ihre Bilder.                                                                   |
| Template-Dateien | CSS- und Template-Dateien, in denen Bildpfade fest hinterlegt sind — etwa Hintergrundbilder.                                                                  |
| Extension-Profil | Drittanbieter-Erweiterungen, für die ein geprüftes Profil vorliegt.                                                                                           |
| Tiefenscan       | Das Auffangnetz: durchsucht alle Datenbanktabellen nach Bildverweisen — auch die von Erweiterungen ohne eigenes Profil.                                       |

Alle sechs lassen sich einzeln abschalten (Optionen → *Scannen*). Das ist selten eine gute Idee: Jeder abgeschaltete Bereich macht die Liste der Kandidaten länger und die Aussage schwächer.

## Der Tiefenscan

Die spezialisierten Bereiche wissen genau, wo sie suchen müssen. Der Tiefenscan weiß das nicht — er geht breit über die Datenbank und sucht nach allem, was wie ein Medienverweis aussieht.

Das macht ihn langsamer, aber unverzichtbar: Er fängt Verweise von Erweiterungen ab, für die es (noch) kein Profil gibt. Wenn Sie ihn abschalten, verlieren Sie genau dieses Sicherheitsnetz.

## Ordner-Referenzen

Manche Verweise zeigen nicht auf eine Datei, sondern auf einen Ordner — etwa ein Zufallsbild-Modul mit dem Parameter „Ordner: `images/banner`“.

Media Organizer wertet das als vollwertige Referenz auf **alle** Dateien darunter, einschließlich Unterordnern. Würde man das übersehen, gälte ein ganzes Bilderverzeichnis als verwaist, nur weil kein Bild einzeln genannt wird.

# Normalisierung: warum Pfade zurechtgebogen werden

Derselbe Verweis kann auf viele Arten geschrieben sein: mit oder ohne führenden Schrägstrich, als absolute URL, mit Joomlas internem `#joomlaImage://`-Zusatz, URL-kodiert. Damit ein Abgleich überhaupt gelingt, führt Media Organizer alles auf eine einheitliche Form zurück.

Ein Sonderfall verdient Beachtung: **Unicode-Normalisierung (NFC)**. Uploads von einem Mac kodieren Umlaute technisch anders als Windows oder Linux — ä kann ein Zeichen sein oder zwei. Ohne Normalisierung würden solche Dateien fälschlich als „ohne Referenz“ gelten. Die Option ist standardmäßig aktiv; lassen Sie sie im Zweifel an.

Über **Bekannte Website-URLs** teilen Sie Media Organizer zusätzlich mit, welche absoluten URLs als lokal gelten — nützlich, wenn Inhalte mit vollständigen Adressen verlinkt sind.

## Ableitungen

Zu vielen Bildern gehören automatisch erzeugte Varianten: Vorschaubilder, verkleinerte Fassungen. Media Organizer erkennt diesen Zusammenhang und vererbt die Einstufung: Ist das Original in Benutzung, gilt die Ableitung ebenfalls als benutzt.

## Was nie erfasst wird

Bestimmte oberste Ordner sind grundsätzlich tabu und werden gar nicht erst inventarisiert — darunter `administrator`, `libraries`, `templates`, `plugins`, `cache` und der Quarantäne-Ordner selbst. Das ist ein Schutz davor, Joomla-Systemdateien anzufassen, selbst wenn ein Medien-Adapter falsch eingestellt ist.

Der Ordner `media/` ist bewusst *nicht* geschützt — dort liegen bei vielen Websites echte Uploads.

## Weiter geht es

- [Erweiterungsprofile und das Coverage-Gate](#)
- [Konfigurationsreferenz](#)

Revision #6

Created 2026-07-30 17:11:08 UTC by Norbert Graup

Updated 2026-08-12 11:56:40 UTC by Norbert Graup