

Was ist Stemming?

Als Stemming (Stammformreduktion) wird in der Sprachwissenschaft und linguistischen Informatik ein Verfahren bezeichnet, das verschiedene morphologische Varianten eines Wortes auf ihren gemeinsamen Wortstamm zurückführt. Durch das Stemming können verbesserte und relevantere Suchergebnisse erzielt werden. Auf jeden Fall führt es zu einer Erweiterung der Ergebnisse.

Verwendung für die SEO

In der Informatik werden zwar seit 1968 verschiedene Stemming-Verfahren getestet, allerdings ist dies bisher vor allem in den USA weit verbreitet. Suchmaschinen nutzen die im Text enthaltenen [Keywords](#), um relevante Suchergebnisse zu erhalten. Dafür setzt Google seit 2004 für die Optimierung der Suchmaschine verschiedene Stemming-Verfahren ein. Diese Veränderungen im Suchalgorithmus ermöglichen eine Differenzierung der Suchanfragen und analysieren so ganze Sätze wesentlich exakter.

Durch das Stemming erfolgt die Identifizierung der Wörter anhand ihres grammatikalischen Wortstamms. Demnach erkennt die Suchmaschine einen Zusammenhang zwischen den Varianten eines Wortes und weist dieser Seite eine Relevanz zu. Wer beispielsweise den Suchbegriff „Suchmaschinenoptimierung“ eingibt, erhält sowohl die Ergebnisse „Optimierung“ als auch „optimieren“. Zusätzlich verringert das Stemming den benötigten Speicherplatz und beschleunigt somit die Suche. Das Stemming reduziert nicht nur die Inhalte auf die Grundformen, sondern optimiert auch ähnliche Begriffe, die dem Themenfeld des Hauptkeywords entsprechen. So könnte zum Beispiel ein Text über Kartoffeln neben der Beschreibung des Gemüses auch über weitere Verwendungsmöglichkeiten und Kartoffelsorten informieren. Auf diese Art würde der Text seine Relevanz zu diesem Begriff erhöhen. Der Nutzer kann das Stemming umgehen, indem er seine gesuchte Begriffskombination in Anführungszeichen setzt.

Das Stemming-Verfahren

Die komplexen Stemming-Algorithmen sind aufgrund der unterschiedlichen sprachlichen Gegebenheiten (Grammatik) auf die jeweiligen Sprachen optimiert. So erkennt der [Algorithmus](#) beispielsweise aus den verschiedenen Wörtern „Hausbau“, „behausen“ sowie „Wohnhaus“ den gemeinsamen Wortstamm „Haus“ und liefert die entsprechenden Suchergebnisse. Für das Stemming stellt es keine Herausforderung mehr dar, ob der Text Wörter im Singular oder Plural enthält oder ob diese durch einen Bindestrich getrennt sind. Das Verfahren ermöglicht sowohl die Reduktion der Wörter auf ihre Grundform als auch die Kürzung der Suffixe oder Präfixe. Die häufigste Variante stellt die Suffix-Entfernung dar.