

Dokumentation zur Verwendung von CSV-Tabellen und UTF-8 Kodierung

Stand: 13. Februar 2023

Die Arbeit mit CSV-Dateien stellt besondere Anforderungen an den Anwendenden. Vor allem die korrekte Darstellung in Excel hängt von einigen Faktoren ab. Sowohl für die Datenqualitätsrichtlinie (DQR) als auch für die Building Master Data Guideline (BMDG) werden CSV-Tabellen verwendet. Eine Besonderheit dieser Tabellen ist, dass die Inhalte durch ein Trennzeichen, in unserem Fall Semikolon, getrennt werden.

In dieser Dokumentation möchten wir Ihnen einen Leitfaden an die Hand geben und Ihnen zeigen, wie Sie diese

- 1 CSV-Dateien korrekt in Excel anzeigen
- 2 CSV-Dateien in Excel bearbeiten
- 3 CSV-Dateien nach der Bearbeitung korrekt kodieren und speichern

können.

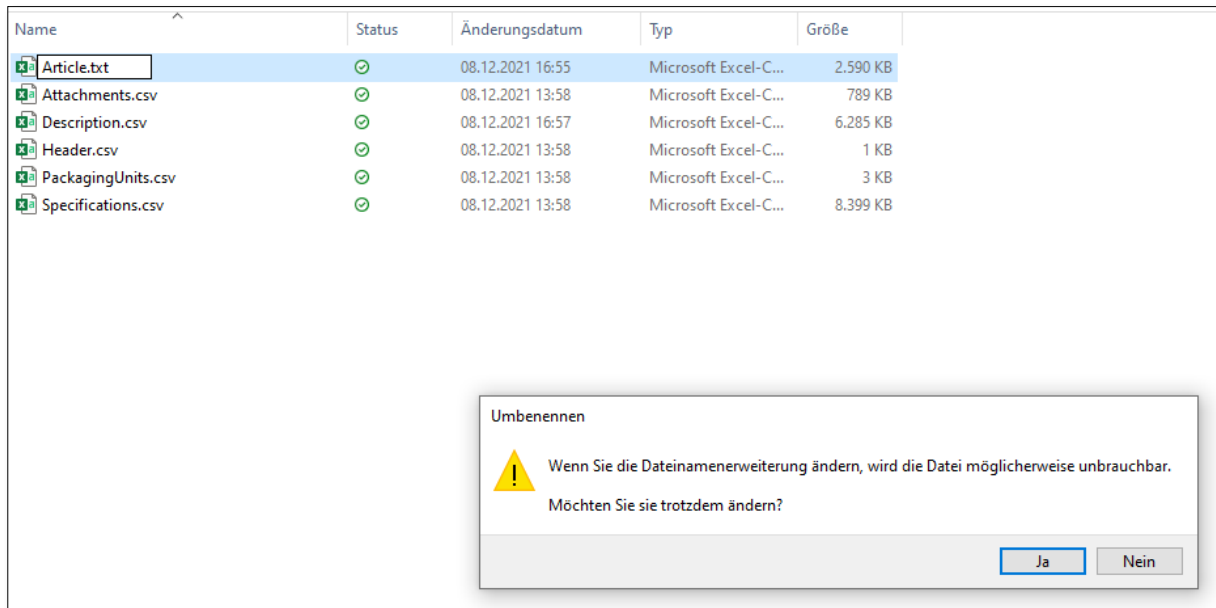
Speziell die Verwendung von Sonderzeichen verschiedener Sprachen führt bei fehlerhafter Kodierung zu Fehlern, die aufwendig behoben werden müssen. Besonderes Augenmerk liegt hierbei auf der Article.csv sowie der Description.csv der BMDG.

Bitte beachten Sie, dass Schritte 2 und 3 unter Umständen nicht nach dieser Anleitung funktionieren, bspw. wenn in Excel das Komma als Trennzeichen voreingestellt ist. Dies kann vor allem im englischsprachigen Raum vorkommen. Wenden Sie sich in diesem Fall bitte an support@arge.de, damit wir gemeinsam mit Ihnen eine Lösung finden können.

Hinweis: Öffnen Sie eine CSV-Datei bitte nicht durch einen Doppelklick. Excel interpretiert die CSV-Datei bestmöglich, jedoch kann dies u.a. zur falschen Interpretation des Trennzeichens, zum Weglassen anführender Nullen (bspw. bei Artikelnummern), dem Verändern des Datumsformats sowie zur Fehlinterpretation der Kodierung führen.

1 CSV-Dateien korrekt in Excel anzeigen

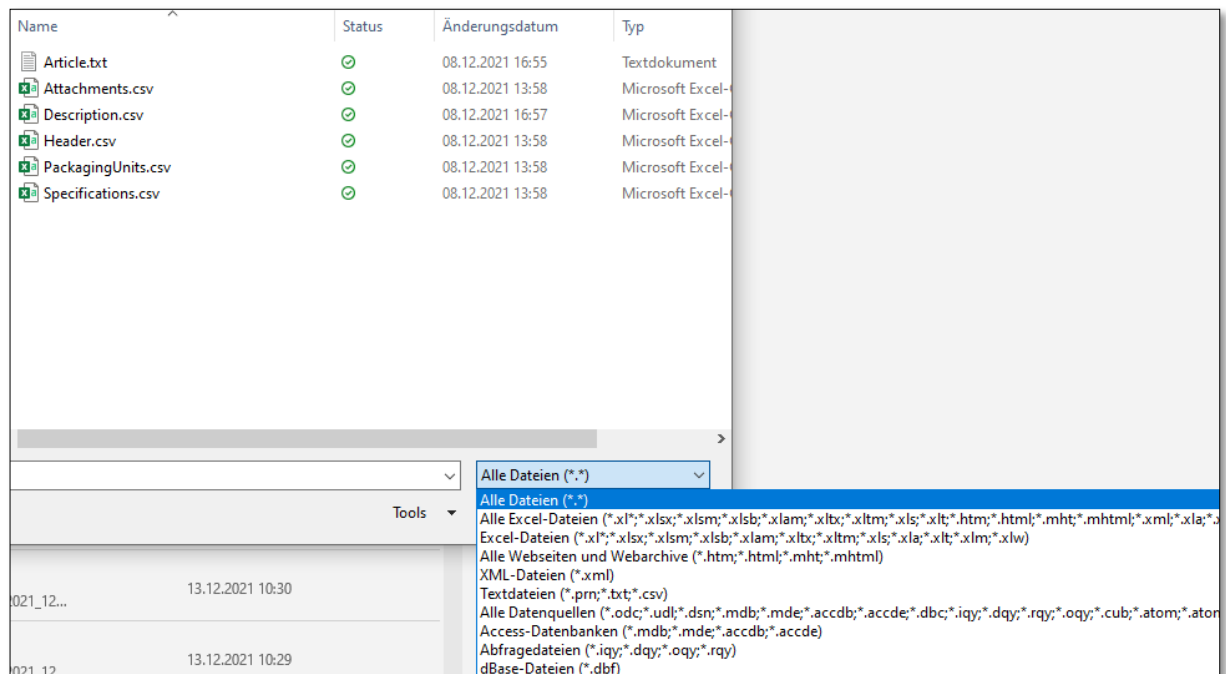
1. Ändern Sie die Dateierweiterung der CSV-Datei, die Sie bearbeiten möchten, in .txt



Bestätigen Sie das Umbenennen mit Ja.

2. Öffnen Sie Excel.

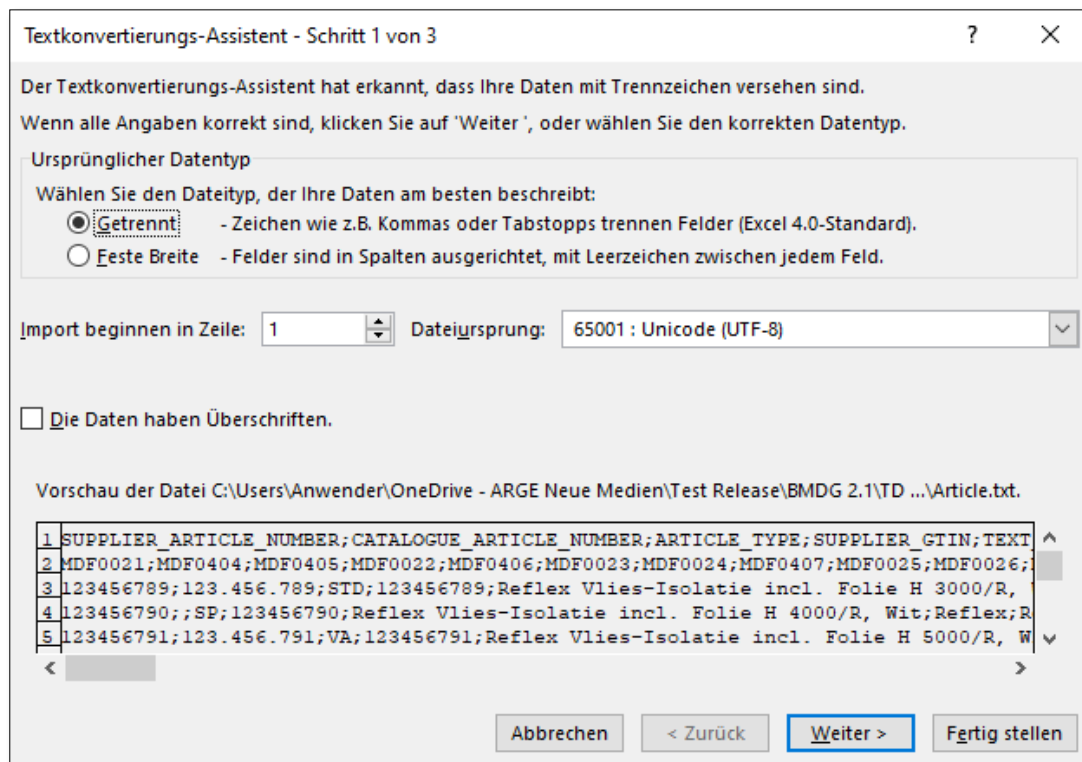
Navigieren Sie über Öffnen → Durchsuchen an den Ablageort der gerade umbenannten Datei. Ändern Sie im Dropdown-Menü in der rechten unteren Ecke die Dateierweiterung von „Alle Excel-Dateien“ zu „Alle Dateien“.



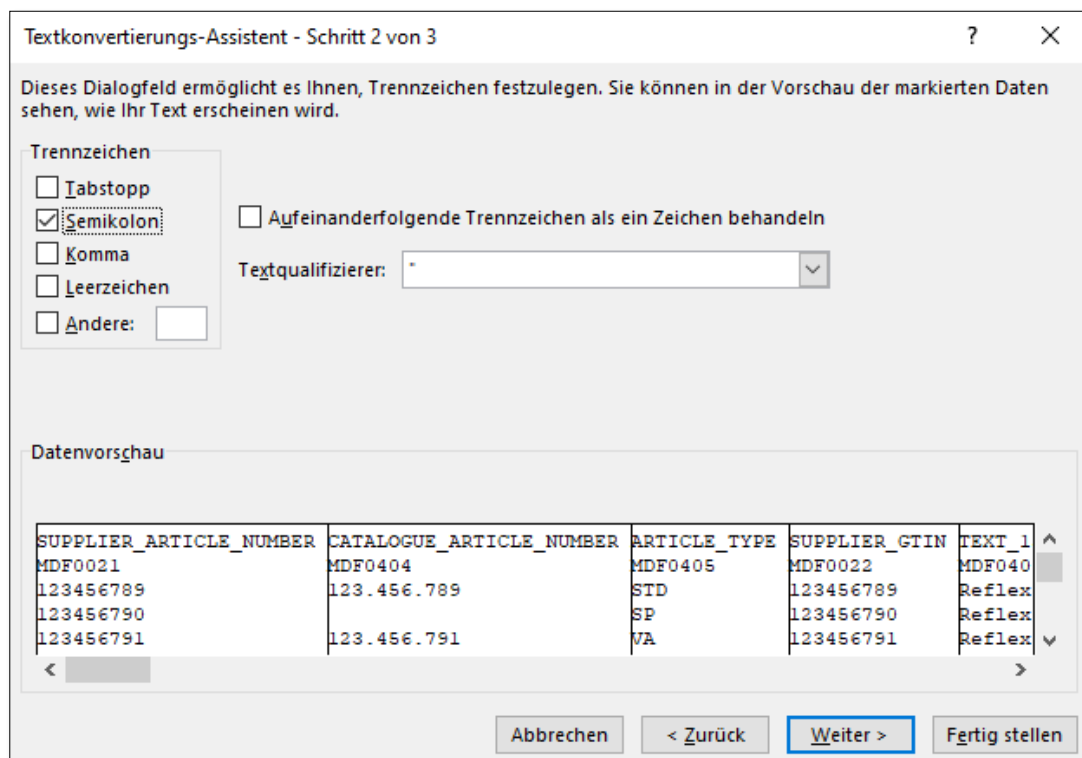
Wählen Sie die Datei aus und klicken Sie auf Öffnen.

3. Nun wird Ihnen der Textkonvertierungs-Assistent angezeigt.

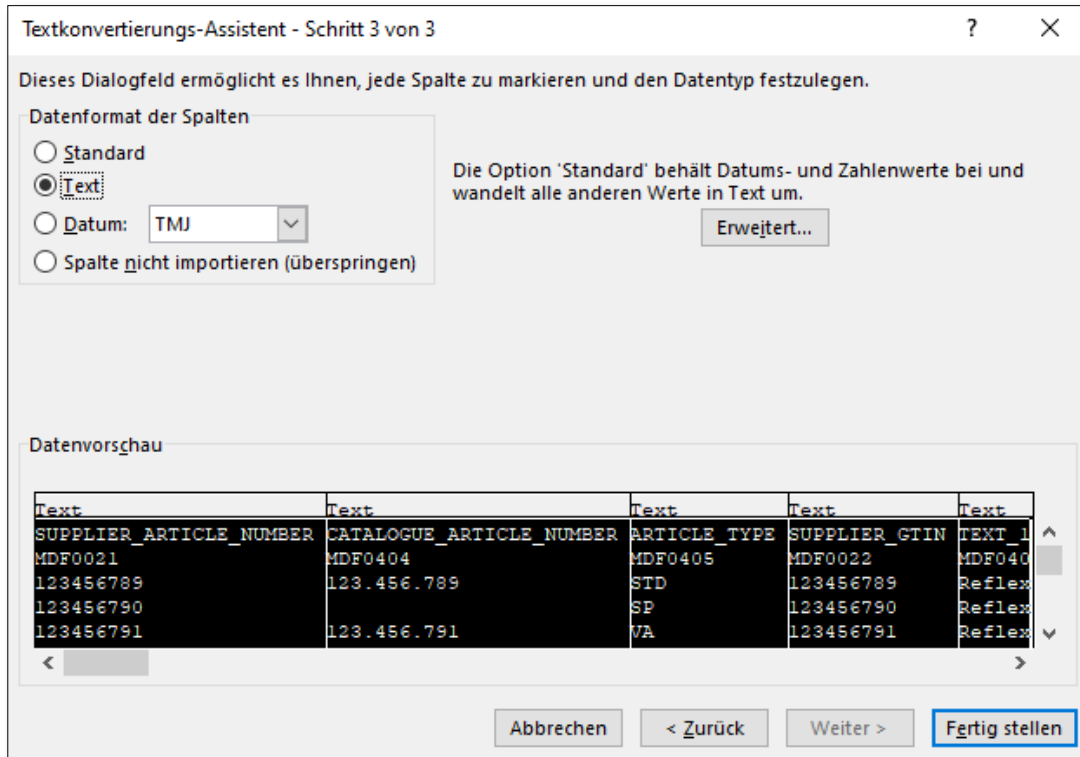
Setzen Sie in Schritt 1 von 3 die Auswahl auf „Getrennt“ und wählen Sie den Dateursprung aus. Im Normalfall sollte dies 65001 : Unicode (UTF-8) sein. Falls Sie nach dieser Auswahl im Vorschauenfenster fehlerhafte Sonderzeichen sehen, wechseln Sie den Dateursprung zu Windows (ANSI) und schauen Sie, ob die Sonderzeichen nun richtig angezeigt werden. Klicken Sie dann auf Weiter.



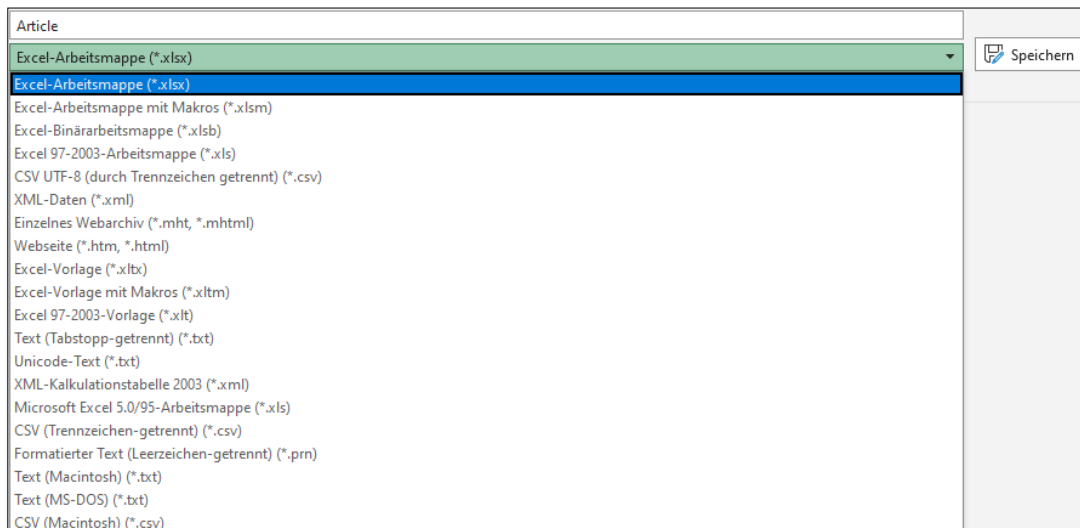
Ändern Sie in Schritt 2 von 3 die Trennzeichenauswahl von Tabstopp zu Semikolon. Klicken Sie dann auf Weiter.



Markieren Sie in Schritt 3 von 3 alle Spalten Ihrer Tabelle. Halten Sie dazu die Shift-Taste gedrückt, ziehen Sie den horizontalen Rollbalken ganz nach rechts und klicken Sie in die letzte Spalte. Ändern Sie nun das Datenformat in Text. Klicken Sie abschließend auf Fertig stellen.



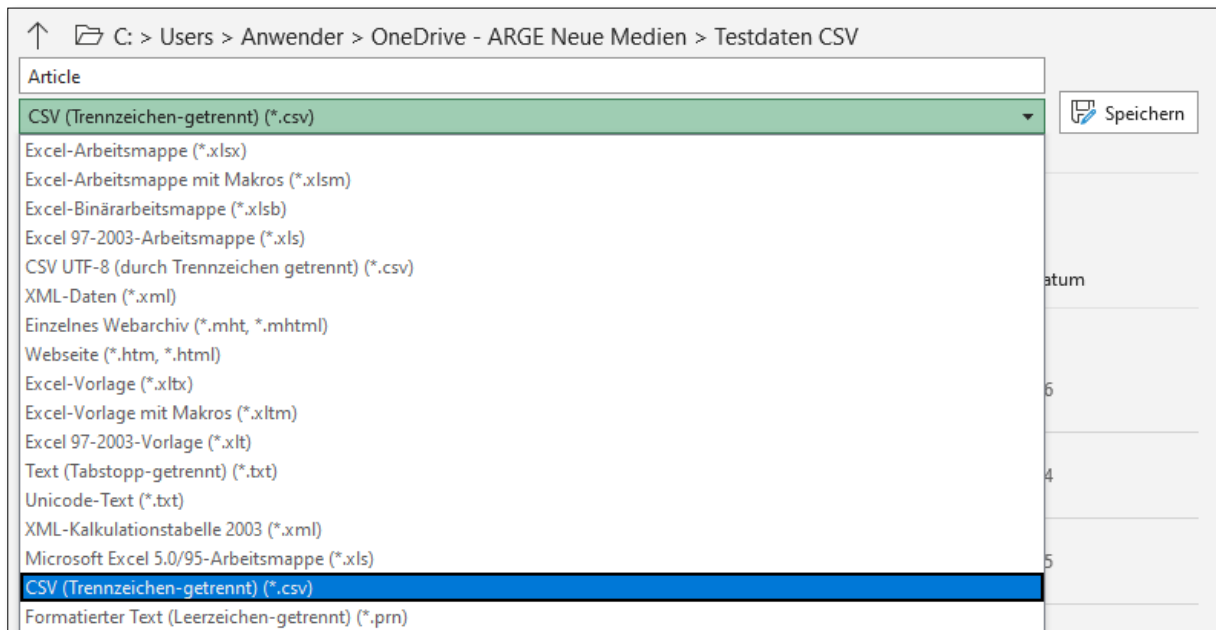
4. Speichern Sie diese Datei nun als xlsx-Datei. Gehen Sie dafür auf Datei → Speichern unter und wählen Sie im Dropdown-Menü Excel-Arbeitsmappe (*.xlsx) aus.



Klicken Sie auf Speichern. Die neue Arbeitsmappe öffnet sich automatisch.

2 CSV-Dateien in Excel bearbeiten

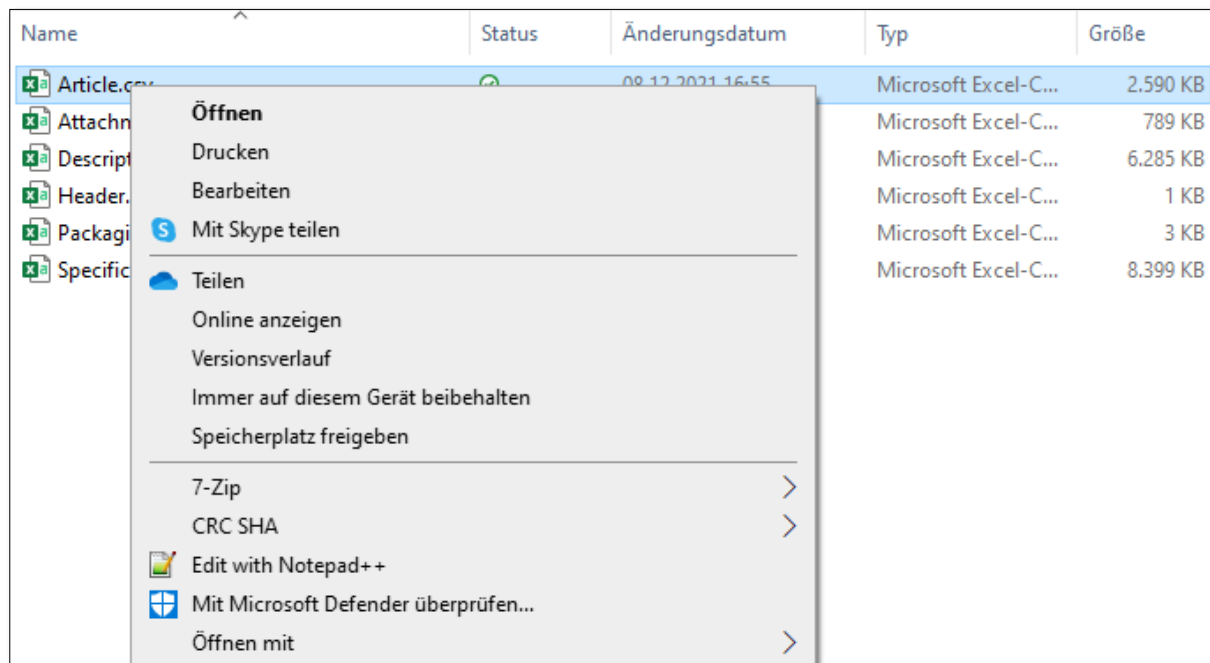
5. Wenn Sie die vorangegangenen Schritte ausgeführt haben, können Sie nun in der erstellten Excel-Datei Ihre Änderungen vornehmen.
6. Wenn Sie alle Änderungen eingetragen haben, speichern Sie die xlsx-Datei. Wenn Sie später weitere Änderungen vornehmen möchten, haben Sie bereits eine Datei, die Sie direkt bearbeiten können.
7. Speichern Sie diese Datei abschließend als CSV-Datei über Datei → Speichern unter und wählen Sie im Dropdown-Menü CSV (Trennzeichen-getrennt) (*.csv). Klicken Sie auf Speichern. Schließen Sie anschließend Excel.



Bitte beachten Sie, dass Excel unter Umständen die Datei nicht mit Semikolon als Trennzeichen speichert. Dies ist vor allem dann der Fall, wenn bspw. wie in Großbritannien das Komma als Trennzeichen in Excel voreingestellt ist. Wenden Sie sich in diesem Fall bitte an support@arge.de, damit wir gemeinsam mit Ihnen eine Lösung finden können.

3 CSV-Dateien nach der Bearbeitung korrekt kodieren und speichern

8. Über einen anderen Editor sollten Sie jetzt die Kodierung überprüfen und ggf. anpassen. Dazu eignet sich bspw. der freie Texteditor Notepad++.
9. Öffnen Sie die gerade erzeugte CSV-Datei über einen Rechtsklick mit der Maus und wählen Sie „Edit with Notepad++“.

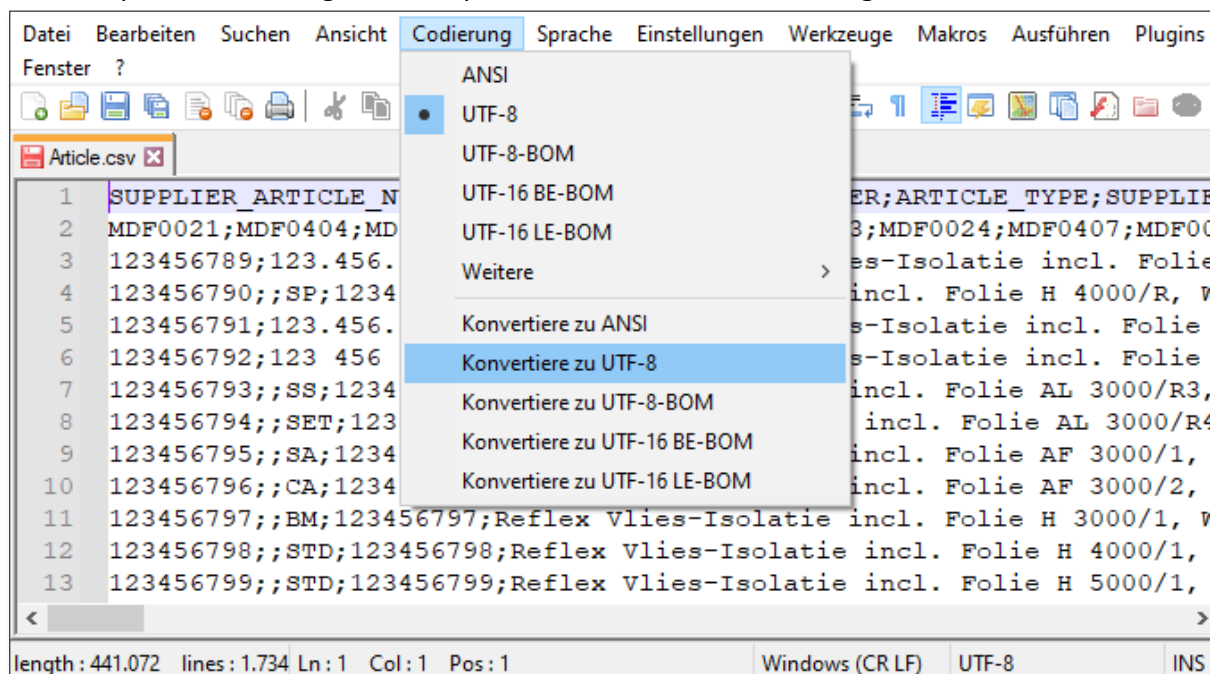


10. Sie sehen nun die CSV-Datei als Textdatei mit Trennzeichen. Wenn Sie ein Semikolon als Trennzeichen sehen, hat Schritt 7 korrekt funktioniert.

11. Jetzt können Sie die Kodierung überprüfen. Sie können die aktuelle Kodierung entweder in der unteren rechten Ecke oder unter Codierung sehen. Wenn die aktuelle Kodierung nicht UTF-8 ist, müssen Sie die Kodierung zu UTF-8 konvertieren.

Klicken Sie dazu auf Codierung → Konvertiere zu UTF-8.

Je nach Spracheinstellung von Notepad++ heißt dies auch Encoding → Convert to UTF-8.



Überprüfen Sie, ob diese Konvertierung erfolgreich und die Sonderzeichen weiterhin korrekt angezeigt werden war. Speichern Sie über das Disketten-Symbol.

Überprüfen Sie die korrekte Konvertierung ggf. durch Wiederholen der Schritte 1 bis 3.