# Szenario A: Neuinstallation auf einem eigenen Server

Beim Aufsetzen einer vollständigen Umgebung und dem Erzeugen eines neuen Modells sind die folgenden Schritte erforderlich.

## Installation der erforderlichen Software und Bibliotheken

* Die Basisinstallation der erforderlichen Bibliotheken (wie *numpy, pandas, scikit-learn, nltk*, etc.) wird im Dokument *AI\_Model\_Dokumentation.md* beschrieben.
* Es wird empfohlen, die im Dokument genannten Bibliotheken und Python-Version auf dem Server zu installieren.

## Datenverarbeitung und Vorbereitung

* Die vorbereitenden Schritte sind in *Datenverarbeitung\_Dokumentation.md* beschrieben. Diese Schritte sollten vor der Modellerstellung durchgeführt werden, um die Rohdaten bereitzustellen und in die korrekte Struktur zu bringen.
* Wesentliche Schritte umfassen die Nutzung der CSV-Dateien und die Bereinigung, etwa mit dem *NullToZero*-Generator (siehe Punkt 1.3).

## Bereinigung der Daten

* Der Prozess des *NullToZero*-Generators (*NullToZero\_Generator.ipynb*, siehe auch *NullToZero\_Generator\_Dokumentation.md*) sorgt dafür, dass alle leeren Zellen in der Matrix korrekt mit Nullen gefüllt werden, sodass ein einheitliches Datenformat für das Modell entsteht.
* Dieser Schritt ist vor der Modellerstellung erforderlich, da er die Qualität der Daten sicherstellt.

## Erstellen des AI-Modells

* Das eigentliche Training und die Erstellung des Modells erfolgt mit dem in *AI\_Model\_Dokumentation.md* beschriebenen Skript *tfidf\_python\_Category.ipynb* (inkl. Kategoriefilter).
* Wenn das Modell erfolgreich trainiert ist, wird es als Datei *tfidf\_model.pkl* gespeichert.

## Modellbereitstellung und Speicherung

* Speichern und Bereitstellen des Modells erfolgen gemäß den Anweisungen in *AI\_Model\_Dokumentation.md*. Das trainierte Modell *tfidf\_model.pkl* ist die Grundlage für Empfehlungen und Analysen.
* Diese Schritte stellen sicher, dass das Modell in produktiven Umgebungen abrufbar und nutzbar ist.

# Szenario B: Update des KI-Modells bei geänderten Daten

Wenn die KI-Umgebung bereits installiert ist, sich aber die Metadaten geändert haben, können die Schritte verkürzt werden, um lediglich das Modell zu aktualisieren.

## Daten bereinigen und aktualisieren

* Führen Sie das *NullToZero\_Generator.ipynb*-Notebook aus, um die neuen Metadaten im vorhandenen Format zu bereinigen.
* Diese Schritte sind kurz in *NullToZero\_Generator\_Dokumentation.md* erläutert.

## Modell neu trainieren

* Das erneute Training des Modells wird anhand der aktualisierten und bereinigten Daten in *AI\_Model\_Dokumentation.md* beschrieben.
* Es ist nur der Modelltrainingsabschnitt erforderlich, Schritte zur Installation und zum Setup entfallen.

## Modell speichern und überschreiben

* Nach dem Training wird das bestehende Modell *tfidf\_model.pkl* durch das aktualisierte Modell überschrieben, ohne dass die Infrastruktur angepasst werden muss.