

# Erkennung/Extraktion

Für die Erkennung/Extraktion von Dokumenten, ist es erforderlich, dass eine Dokumentenklasse für jedes Dokument zuvor bestimmt wurde.

Dies kann über die Importparameter erfolgen oder über die Klassifikation von SQUEEZE.

Die Bestimmung der Dokumentenklasse ist daher so wichtig, da die Dokumentenklasse den Feldkatalog bestimmt. Dieser wiederum legt fest, welche Werte auf dem Dokument mit welchem Verfahren (Lokatoren) gesucht werden sollen. Die ermittelten Werte werden entsprechend der Felddefinition formatiert.

Über den Feldkatalog wird außerdem bestimmt welche der Felder Pflichtfelder sind und ob der Wert eines Feldes durch einen Anwender bestätigt werden muss.

Da sich die Klassifikation des Systems auch mal irren kann, ist es natürlich während der Validierung der Dokumente möglich, die Dokumentenklasse anzupassen. Eine Änderung der Dokumentenklasse führt zur erneuten Erkennung des Dokuments mit dem Feldkatalog der neuen Dokumentenklasse.

---

Revision #2

Created 2020-03-04 20:14:33 UTC by Phillip Langer

Updated 2020-03-10 11:49:59 UTC by Jasmin Ruß