Bachelorarbeit Manuel Lang

* **Ausgangssituation**

Die APLICONUS GmbH spezialisiert sich auf ergonomische und zugleich leistungsstarke Geschäftsanwendungen für PCs, Windows Tablets und Windows Phones.

Eine dieser Anwendung ist APLICONUS infield, ein Personal Information Manager. Damit können Mitarbeiter Kontakte, Reisen, Zeiterfassung und Bewirtungen erfassen und verwalten.

Dabei kann ein Benutzer bspw. Kosten, die er erstattet haben möchte, eintragen. Aktuell muss der Benutzer diese von Hand eintragen.

Ein Premium-Abonnement gegenüber einer geringfügigen monatlichen Pauschale soll ermöglichen, Belege einzuscannen und diese somit in das System zu importieren.

* **Zielsetzung**
  + Implementierung einer Bibliothek, um Reisekostenbelege zu erkennen, mit Hilfe von optischer Zeichenerkennung auszulesen und mit Cloud-basiertem Maschinellen Lernen platformunabhängig zu verarbeiten
* **Aufgaben**
  + Zeichenerkennung mittels Windows OCR (Optical Character Recognition)
    - Ermitteln von verschiedenen Daten, z.B.
      * Aussteller des Belegs (Name, Telefonnummer)
      * Gesamtbetrag (Brutto)
      * Mehrwertsteuer
      * Rechnungsdatum
      * Leistungszeitraum
  + Clustering (Support vector machines) mittels Azure ML (Rohdaten -> Nutzdaten)
    - Kategorie ermitteln (Supermarkt, Hotel, Reise, Bewirtung,…)
    - Gewünschte Felder für gefundene Kategorie ermitteln
  + Erkannte Zeichen auf Sinnhaftigkeit überprüfen (lokal)
    - Vergleich mit Wörterbüchern
    - Straßennamen (Open Map)
    - Barcodes
* **Abgrenzung**
  + Verwaltung von Dokumenten/Zahlungsbelegen
    - Übersicht über bereits verwendete Belege
    - Bearbeiten der Belege
    - Hinzufügen von neuen Belegen
    - Löschen von Belegen
    - Erstellen von Anmerkungen
  + Einbetten der gewonnen Daten in APLICONUS infield