

APP zum Scannen, Archivieren und Suchen

SWP (inkl. Bachelorarbeit) – Peter Schartner – V01 – 2021-09-06

Kurzfassung:

Zu entwickeln ist eine Android- bzw. iOS-App, die das einfache Einscannen von Papierdokumenten und das sichere Archivieren und effiziente Suchen der elektronischen Kopien unterstützt.

Team: 2 Personen (Informatik und/oder Informationsmanagement)

Vorausgesetzte Kenntnisse:

- Programmiererfahrung mit Java
- Grundkenntnisse IT-Sicherheit
- Kreativität (GUI-Design)
- Android- bzw. iOS-Kenntnisse sind von Vorteil aber kein Muss

Plattform: Android (und ggf. iOS)

Bearbeitungsbeginn: Ab sofort

Betreuer: Peter Schartner (peter.schartner@aau.at)

Grobspezifikation der gewünschten Funktionalität:

- 1) Scannen inkl. Entzerrung (vgl. MS OfficeLens)
- 2) Beschlagwortung
- 3) Suche
- 4) Accountverwaltung und Verschlüsselung
- 5) Mögliche Erweiterungen

Scannen

Zunächst einfaches frontales Abfotografieren (mit manueller Beschneidung) ohne Entzerrung. In den Metadaten wird zumindest das Datum des Scanprozesses erfasst. Optional kann ein Datum und ein Titel des Dokuments erfasst werden, die dann in den Dateinamen einfließen (JJJJ-MM-TT Titel). Diese beiden Metadaten sollen auch nachträglich noch vergeben werden können. Ablage der Dokumente als JPG und/oder PDF und Ablage der Metadaten sowohl in den Dokumenten (als JPG-EXIF-Daten bzw. PDF-Metadaten) und als getrennte strukturierte Datei (txt/xml).

In weiterer Folge soll das eingescannte Dokument durch Ausnutzung von Kontrast und/oder Farbe automatisch entzerrt werden. Wenn keine automatische Entzerrung möglich ist, dann soll dies manuell durchgeführt werden können.

In erster Linie soll die App lokal laufen, aber in einer weiteren Ausbaustufe sollen Daten von unterschiedlichen Geräten (z.B. Handy und Tablet) in einem Account (mit Cloud-Repository) zusammengeführt werden können.

Beschlagwortung

Umsetzung einer einfach nutzbaren individuellen Beschlagwortung, wobei die hierarchisch strukturierten Schlagwörter nach Kategorien und Unterkategorien vorgegeben werden und erweiterbar sein sollen. Erweiterungen erfolgen entweder global über die Einstellungen der App, oder beim Erfassen eines

bestimmten Dokuments. Wird ein neues Schlagwort hinzugefügt, sollen ähnliche bereits vorhandene Schlagwörter angezeigt werden (String-Similarity).

Beispiel 1: Dokumente im Uni-Umfeld

- Kategorien: Lehre, Forschung, Administration, Privat
- Unterkategorien Lehre: WS20, SS21, BA-Arbeiten, MA-Arbeiten, ..., A&D, ETI, SS, PV, ...
- Schlagwörter unter WS20: ETI, PV, ...
- Schlagwörter unter ETI: WS20, WS21, ...

Beispiel 2: Rezepte (könnte auch eine Unterkategorie von „Privat“ sein)

- Kategorie: Frühstück, Vorspeise, Suppe, Hauptspeise, Nachspeise
- Unterkategorie: kalt, warm
- Schlagwörter: Zubereitungsmethoden und Hauptzutaten

Auf Ebene der Dokumente soll eine Bewertung (z.B. 1-5 Sterne) vergeben werden und Kommentare zum Dokument eingegeben werden können.

Suche

Die Suche basiert im Wesentlichen auf der Hierarchie von Kategorien, Unterkategorien und Schlagwörtern. Diese werden sukzessive ausgewählt und so wird die Menge an verfügbaren Dokumenten eingegrenzt. Zunächst wird deren Anzahl angezeigt und wenn die Anzahl überschaubar ist, die Namen der Dokumente. Die Anzeige erfolgt alphabetisch, chronologisch oder nach der Bewertung sortiert.

Alternativ zur geführten Suche sollen Kategorien und Unterkategorien dynamisch als Ordnerstruktur angezeigt werden, sodass man auch über diesen Weg zu den Dokumenten navigieren kann. Ein Dokument mit Schlagwörtern „Lehre“, „ETI“ und „WS20“ soll bspw. über die Pfade „Lehre -> ETI -> WS20“ und „Lehre -> WS20 -> ETI“ auffindbar sein.

Optional: Accountverwaltung und Verschlüsselung

Optional soll eine App unterschiedliche Accounts anlegen und verwalten können (z.B. Trennung zwischen „beruflich“ und „privat“, unterschiedlichen Bereichen oder Personen).

Werden Daten in einer Cloud gespeichert ist State-of-the-Art-Verschlüsselung verpflichtend, bei lokaler Speicherung ist die Verschlüsselung per Default aktiviert, sie kann aber deaktiviert werden. Ebenso kann die PIN zum Schutz der App (und der Schlüssel) bei lokaler Nutzung deaktiviert werden.

Die Schlüssel zur Verschlüsselung der Daten müssen (unter Nutzung von Dateiname und Randomizer) von einem zufälligen Masterschlüssel abgeleitet werden. Dieser Masterschlüssel und die Randomizer sind mit einem aus der PIN abgeleiteten Schlüssel zu schützen. Der Masterkey darf nicht direkt aus der PIN abgeleitet werden!

Mögliche Erweiterungen

- Ersetzen, Umstrukturierung und Zusammenführung von (Unter-)Kategorien und Schlagwörtern.
- Sprachbasierte Suche
- Versenden von einzelnen Dokumenten und ganzen Ordnern per Mail
- Import von Dokumenten (jpg, PDF, ...) und ganzen Ordnern
- Umwandlung von eingescannten Dokumenten (jpg) in PDF
- Eigene Ideen sind willkommen!