
PRAKTIKUMSBERICHT

Herr
Hannes Steiner

Entwicklung eines Dokumenten Scanners

**Digitalisierung der Verwaltung an der Hochschule
Mittweida**

2020

PRAKTIKUMSBERICHT

Entwicklung eines Dokumenten Scanners

**Digitalisierung der Verwaltung an der Hochschule
Mittweida**

Autor:
Hannes Steiner

Studiengang:
Softwareentwicklung

Seminargruppe:
IF17wS-B

Erstprüfer:
Prof. Dr. Mark Ritter

Zweitprüfer:
N.N.

Mittweida, März 2020

Bibliografische Angaben

Steiner, Hannes: Entwicklung eines Dokumenten Scanners, Digitalisierung der Verwaltung an der Hochschule Mittweida, 21 Seiten, Hochschule Mittweida, University of Applied Sciences, Fakultät Angewandte Computer- und Biowissenschaften

Praktikumsbericht, 2020

Dieses Werk ist urheberrechtlich geschützt.

I. Inhaltsverzeichnis

Inhaltsverzeichnis	I
Abbildungsverzeichnis	II
Tabellenverzeichnis	III
Abkürzungsverzeichnis	IV
1 Einleitung	1
2 Problemstellung	3
2.1 Digitalisieren der Klausur-Daten	3
2.2 Klausuren-Vorlage	3
2.3 Vorlage verbessern	4
2.4 Weitere Anmerkungen	4
3 Konzept des Documenten-Scanners	5
4 Anforderungen	7
5 Entwicklung des ersten Prototyps	9
6 Weitere Entwicklung	11
6.1 App Architektur	11
7 Grenzen der App	13
7.1 Probleme beim Erkennen von Dokumenten	13
7.2 Probleme der Klausur-Vorlage beim Scannen	13
7.3 Weitere	13
8 Templates	15
9 Ausblick	17
Literaturverzeichnis	19

II. Abbildungsverzeichnis

III. Tabellenverzeichnis

IV. Abkürzungsverzeichnis

bzw. beziehungsweise, Seite 3

Fakultät CB Fakultät für angewandte Computer- und Biowissenschaften, Seite 3

OCR optical character recognition, deutsch: optische Zeichen Erkennung, Seite 5

1 Einleitung

Digitalisierung wird gängig als Integration von digitaler Technologie in den Alltag verstanden, und soll helfen Zeit einzusparen [1]. Mit diesem Gedanken initiierten die Mitarbeiter Holger Langner und Falk Schmitsberger der Hochschule Mittweida, das Projekt *Memo Space*. Im Zuge dessen sollen kleinere Forschungsergebnisse entstehen, die richtungsweisend für die Digitalisierung der Verwaltung von Lehr- und Forschungseinrichtung sind.

Eine der ersten Ideen ist es, die Arbeit von Klausur-Prüfern zu erleichtern. Diese müssen, nachdem die Klausuren kontrolliert wurden, die Benotungen, sowie die Eckdaten der Studenten, in ein digitales Format bringen. Grund dafür ist, dass die Noten in das Notensystem der Einrichtung eingetragen werden müssen.

Im Rahmen eines Forschungspraktikums an der Hochschule Mittweida arbeiteten der Student Tobias Kallauke und der Verfasser, gemeinsam an einer Lösung zur Digitalisierung dieses Arbeitsschrittes. Sie entwickelten... (Klausuren-Vorlage, App, Server, OCR? evtl.) / Dabei entstanden .../ Der Verfasser programmierte...

Was noch?

2 Problemstellung

Hochschulmitarbeiter sitzen zum Ende eines Semesters über Tage an der Kontrolle von Klausuren. Diese Aufgabe muss stets mit hoher Konzentration erledigt werden, und lässt sich aber in den meisten Fällen nur schwer durch Maschinen ersetzen. Unter keinen Umständen dürfen bei der Bewertung Fehler vorkommen, was jedoch bei der kognitiven Last der Prüfer immer wieder passiert. Auch nach der Durchsicht der Prüfungsaufgaben ist eine hohe Achtsamkeit wichtig. Denn anschließend wird die Benotung in eine digitale Tabelle überführt. In diese muss die Matrikelnummer, der Vor- und Nachname, sowie die Note des Studenten eingetragen werden. Hier kommt es vor allem bei der Matrikelnummer und der Zensur auf die Richtigkeit jedes Zeichens drauf an.

2.1 Digitalisieren der Klausur-Daten

Für genau diesen Vorgang des Digitalisierens wird eine Lösung gesucht. Die Prüfer sollen so bequem und möglichst zeitsparend diese Aufgabe verrichten, ohne dabei ihre Aufmerksamkeitsspanne zu überlasten. Des Weiteren müssen die Ergebnisse der Prüfungen, sowie die Eckdaten der Studenten in ein geeignetes digitales Format gebracht werden, um es der Notenfreigabe weiterzuleiten. Darüber hinaus empfiehlt es sich, digitale Kopien der Klausuren abzuspeichern, da/um ... (warum genau, sollen die gespeichert werden? meine/unsere Idee Online Klausureneinsicht -> ins Fazit/Ausblick (hat viele "Probleme"))

2.2 Klausuren-Vorlage

- ist es eher ein Gestaltungsleitfaden "style guide" oder eine Klausuren-/Layout-Vorlage?

Eine Klausuren-Vorlage bzw. ein Gestaltungsleitfaden für Klausuren der Fakultät *Angeordnete Computer- und Biowissenschaften* (Fakultät CB) bietet außerdem die Möglichkeit der Kontrolle des Prüfers an. Genauer ist es durch das vorgegebene Layout der Klausur möglich, einen Teil der Arbeit des Prüfenden auf Fehler zu untersuchen. Auf dem Deckblatt der Klausur ist eine Tabelle, mit drei Zeilen und für jede Aufgabe eine Spalte. In der ersten Zeile befinden sich die Nummern der Aufgaben. In der zweiten, die zu erreichenden Punkte der Aufgabe. Und in der dritten Zeile trägt der Prüfer die erbrachten Punkte des Studenten ein. Unter der Tabelle befindet sich ein Feld für die erreichte Gesamtpunktzahl, sowie ein Feld für die aus den Punkten resultierende Note. Die Punkte pro Aufgabe, die Gesamtpunktzahl und die Zensur stehen in (Korrelation?) Relation zu einander, sodass aus den Punkten der Aufgaben die beiden anderen Felder errechnet werden und mit den Ergebnissen des Prüfers abgeglichen werden könnten.

Eine weiteres Merkmal der Klausuren-Vorlage ist ein Feld, für die vom Studenten erreichten Punkte über jeder Aufgabenstellung. Die dortige Angabe sollte mit der, in der Tabelle auf dem Deckblatt übereinstimmen und bietet somit noch eine weitere Möglichkeit der Kontrolle an.

2.3 Vorlage verbessern

Templates entwerfen, die optimiert für die Digitalisierungs-Prozesse sind ...

2.4 Weitere Anmerkungen

TODO:

Ferner soll bei der Lösung von der Anschaffung neuer Technologie und Geräte abgesehen werden. Grund dafür sind neben den Anschaffungskosten, die Idee, dass das Ergebnis des Forschungsprojekts in weiteren Lehr- und Forschungseinrichtung Anwendung finden sollte.

Was noch?

Soll hier schon rein, dass eine App, ein Webportal, etc. entstehen soll? Oder soll die Grund Idee zur Lösung in ein extra Kapitel?

3 Konzept des Documenten-Scanners

Aus der Problemstellung heraus entstand die Idee einer Dokumenten-Scanner App. Diese soll die wichtigen Daten auf dem Deckblatt einer Klausur, wie Vor- und Nachname des Studenten, seine Matrikelnummer sowie die Note erkennen, digitalisieren und in ein für die Notenfreigabe geeignetes Format bringen. Die digitalisierten Daten sollen anschließend an einen Server gesendet werden, wo sie und die beim Einscannen entstandenen Bilder gespeichert werden. Bei Verwendung der Klausuren-Vorlage der Fakultät CB oder ähnlichen, die eine Punkteübersicht haben, soll es außerdem möglich sein, die Punkte sowie die Note auf Richtigkeit zu überprüfen.

Zur Digitalisierung und Kontrolle der Daten werden Scan-Vorlagen verwendet. Diese muss vor dem Einscannen der ausgefüllten Klausuren erstellt werden, um zu gewährleisten, dass die richtigen Daten digitalisiert werden. Beim Erstellen einer Scan-Vorlage geht man wie folgt vor:

- Zuerst wird jede Seite der Klausur fotografiert. Dabei wird in jedem Bild das Dokument erkannt, vom Hintergrund getrennt bzw. ausgeschnitten und perfekt ausgerichtet. (Kontrast erhöht -> Schrift besser zu erkennen, ...)
- Anschließend markiert man diejenigen Regionen auf jeder Seite, die zu digitalisieren und/oder kontrollieren sind.
- Des Weiteren kann man aus allen markierten Regionen noch diese auswählen, die in Relation stehen. Diesen Relationen werden danach noch weitere Eigenschaften zugeteilt, um sie später bei der Kontrolle richtig zu analysieren. So eine Eigenschaft könnte sein, dass es sich um einen Vergleich zwischen zwei in Relation stehenden Regionen handelt. Z.B. ist die eine Region die Zelle in der Punkteübersichts-Tabelle, in der die erreichten Punkte zu Aufgabe 1 rein geschrieben werden sollen und die zweite Region die Stelle über der Aufgabenstellung von Aufgabe 1, in der ebenfalls die erreichten Punkte eingetragen werden sollen. Weitere Beispiele für solche Relationen und Eigenschaften sind die Summanden für die erreichte Gesamtpunktzahl und die Summe selbst, oder auch noch die daraus resultierende Note.
- Als letztes speichert man die Vorlage ab, welche dann automatisch an einen Server gesendet wird, so dass auch andere diese Vorlage nutzen können.

Nach der Erstellung einer Vorlage kann das Dokument digitalisiert werden. Dazu scannt man die Seiten in der selben Reihenfolge, wie in der Vorlage ein. Die App digitalisiert mit Hilfe der Vorlage und optical character recognition (OCR) den Inhalt in den vordefinierten Regionen. Im Anschluss an die Texterkennung kann das Resultat nochmal überprüft und korrigiert werden, bevor es an den Server zum Abspeichern gesendet wird.

... noch mehr? (Auf Tobias warten...)

Ein wichtiger Vorteil einer App gegenüber anderen Lösungen ist, dass so gut wie jeder Mitarbeiter an einer Lehr- und Forschungseinrichtung ein eigenes oder Zugang zu einem Smartphone oder Tablet hat.

- Einspeisung von erwarteten Ergebnissen noch mit rein nehmen? (Noten: ,0 ,3 oder ,7; Seminargruppen und Klausur Teilnehmer.)
- Noch mehr zu OCR sagen, was es ist, wie es geht, komplettes eigenes Kapitel?

4 Anforderungen

- ios app:
- ios 13.0 > version, wegen swiftui und vision / vision kit -> bedeutet geräte ab dem 6s, aber nicht ältere (Prozessor ist entscheidend)
- auch iPads sind recht einfach möglich dank SwiftUI!!!
- vlt ein paar worte zu swift, swiftui, vision und visionkit ...
- dokumente ein scannen
- vorlagen erstellen, um wichtige Bereiche zu erfassen
- mit der vorlage dokumente einscannen
- auf dem gerät oder auf dem server die texte erkennen
- vorlagen auf einem server speichern und abrufen
- ergebnisse an den server schicken, nicht auf der App speichern
- Datenschutz/sicherheit beachten, auch bei Server und APIs
-

5 Entwicklung des ersten Prototyps

- beschreiben wie die einzelnen Views/Seiten nun aussehen und funktionieren? oder eher den Prozess der Entwicklung?
- Bild mit Wireframe aller Views und deren Workflow?
- AppState erklären? (Single source of truth) -> später AppStore mit mehreren States
- erklären deskew / Dokument ausrichten?
- Umrechnung der Bilder vom Template aus und zuschneiden der neuen Regionen erklären
- Attribute hinzufügen (den Vorgang) -> Rechteck einzeichnen...
- Verwendete Programme, Sprache, etc.
- ein paar Worte, wie gut die Entwicklung lief (Simulator vs. echtes Gerät), wo sind die Grenzen des Simulators (Keine Kamera), wo waren Probleme mit dem echten Gerät...

6 Weitere Entwicklung

6.1 App Architektur

- redux like app store mit states erklären? (single source of truth)
- auf die API eingehen? oder nur tobias?
-

7 Grenzen der App

7.1 Probleme beim Erkennen von Dokumenten

- Gleich-farbiger Hintergrund
- Wenig Licht
- Hintergrund mit starken Kanten
- Runde Ecken, keine Ecken
- Starke Kanten im Bild (schwarzer Kreditkartenstreifen)
- zu jedem möglichem vlt. dann ein Beispiel Bild

7.2 Probleme der Klausur-Vorlage beim Scannen

- Schrift zu klein
- zu wenig Platz
- zu viel Platz
- Überschneidungen
- ...

7.3 Weitere ...

- ...

8 Templates

- Template und den Entwicklungsprozess vorstellen,
- wieso weshalb warum muss das nun so aussehen?
- Was kann an der neuen Vorlage immer noch optimiert werden?
- Welche Probleme konnten behoben werden?

9 Ausblick

- Es müssen nicht nur Klausuren sein, sondern alles mögliche (Krankenscheine, Urlaubsscheine, was auch immer, nur DB muss angepasst werden. (oder automatisches generieren von DB-Tabellen an Hand der Template Attribute))
- Server mit Bildern als Klausuren-Einsicht nutzen -> allerdings viele Probleme (Klausuren würden kopiert werden -> Profs mehr Arbeit, keinen direkten Kontakt zum Prof wegen Fragen, Verbesserungen oder Anmerkungen, ...)
- QR-Code-Idee um Vorlagen weg zu lassen -> Programm oder Plugin (Word/LaTex) was die QR-Codes dann automatisch erstellt und richtig einfügt. (Wichtige Daten sind da hinterlegt, Vor- und Nachteile von QR-Codes, ...)
- Ist die App nun schneller als der normale Umgang bleibt offen -> Bachelorarbeit knüpft da an...

Literaturverzeichnis

[1] Michael Graf, Partner bei PwC

<http://www-cs-faculty.stanford.edu/~uno/abcde.html>

Erklärung

Hiermit erkläre ich, dass ich meine Arbeit selbstständig verfasst, keine anderen als die angegebenen Quellen und Hilfsmittel benutzt und die Arbeit noch nicht anderweitig für Prüfungszwecke vorgelegt habe.

Stellen, die wörtlich oder sinngemäß aus Quellen entnommen wurden, sind als solche kenntlich gemacht.

Mittweida, 13.03.2020