



**Zavod za elektroniku,
mikroelektroniku, računalne
i inteligentne sustave**

Programsko inženjerstvo

Igor Stančin

Digitalizacija

Zagreb, listopad 2021.

Potrebno je izraditi web aplikaciju koja će ubrzati digitalizaciju računovodstvenim tvrtkama. Glavna funkcionalnost aplikacije je detekcija dokumenta na slikama i izrada OCR-a (optical character recognition) detektiranog teksta. Mora postojati mogućnost učitavanja do 50 slika istovremeno. Radi jednostavnosti detekcije dokumenta, može se pretpostaviti da će pozadina uvijek biti drugačije boje od dokumenta. Kako bi se dodatno osiguralo da kvaliteta detektiranog dokumenta bude dovoljno dobra za OCR, za svaki učitani dokument potrebno je provjeravati je li kut slikanja bio ispravan, odnosno potrebno je utvrditi da je detektirani dokument približno pravokutnog oblika. Nakon što je napravljen OCR dokumenta, korisniku se prikazuje sažetak dokumenta. Korisnik može skenirani dokument označiti kao točno skenirani ili kao krivo skenirani. Što god korisnik odabrao, dokument i korisnikov odabir spremaju se u bazu.

Korisnici aplikacije su zaposlenik, revizor, računovođa i direktor. Svi korisnici imaju mogućnost skeniranja dokumenata. Zaposlenik može aplikacijom skenirati dokumente i može vidjeti povijest svih svojih skeniranja. Dokument koji je skeniran i potvrđen kao točan od strane zaposlenika šalje se revizoru. Revizor provjerava sve pristigle dokumente i svaki preusmjerava računovođu koji je zadužen za taj tip dokumenata. Ako revizor skenira dokumente aplikacija će automatski iz dobivenog teksta odrediti kojem računovođi se šalje dokument.

Računovođa dobivene dokumente arhivira. Aplikacija prilikom arhiviranja dokumentu dodjeljuje jedinstveni broj arhiva. Također, računovođa ima opciju slanja dokumenta direktoru na potpis prije arhiviranja. U tom slučaju direktor dobiva obavijest, a nakon potpisa dokumenta šalje se obavijest računovođi da sada može arhivirati traženi dokument. Direktor može vidjeti povijest svih dokumenta te povijest i statistike svih zaposlenika.

Postoje tri tipa dokumenata – računi, ponude i interni dokumenti. Računi će u svom tekstu nakon OCR-a imati oznaku računa koja je veliko slovo R te šest znamenaka, oznaka ponude će imati veliko slovo P i devet znamenaka, a oznaka internog dokumenta „INT“ i četiri znamenke. Računi osim oznake sadrže ime klijenta, artikle s cijenama i ukupnu cijenu. Ponude su kao računi, ali ne sadrže ime klijenta. Interni dokumenti sadrže samo nestrukturirani tekst.

Detaljnija pojašnjenja moguća su na laboratorijskim vježbama kod demonstratorice Klara Marijan (klara.marijan@fer.hr) i asistenta Igor Stančin (igor.stancin@fer.hr) prema važećem rasporedu objavljenom na stranicama predmeta. Svaki dolazak je potrebno **najaviti mailom barem jedan dan ranije.**