



POLITECHNIKA POZNAŃSKA

WYDZIAŁ INFORMATYKI I TELEKOMUNIKACJI
Instytut Informatyki

Projekt - Hurtownie Danych i Przetwarzanie Analityczne

PARSER DOKUMENTÓW PDF I SKŁADNICA DANYCH MEDYCZNYCH

Bartosz Paulewicz, bartosz.paulewicz@student.put.poznan.pl

Jakub Frieske, jakub.frieske@student.put.poznan.pl

Bartosz Kowalski, bartosz.t.kowalski@student.put.poznan.pl

Wojciech Marciniak, wojciech.to.marciniak@student.put.poznan.pl

POZNAŃ 2022

Spis treści

1	Wstęp	9
1.1	Cel i zakres pracy	9
2	Analiza wymagań	11
2.1	Wymagania funkcjonalne	11
2.2	Wymagania pozafunkcjonalne	12
2.3	Opis dokumentu badania ALAB	12
2.3.1	Zawartość dokumentu ALAB	12
2.3.2	Opis głównych sekcji	14
	Nagłówek	14
	Sekcja dotycząca badania	14
	Sekcja końcowa	14
3	Opis implementacji	17
3.1	Schemat bazy danych	17
3.1.1	Spis tabel	17
3.1.2	Pacjenci	18
3.1.3	Dokumenty	18
3.1.4	Dokumenty ALAB	18
3.1.5	Wykonali zatwierdzili	19
3.1.6	Badania	19
3.1.7	Wyniki badań	19
3.1.8	Wartości	19
3.1.9	Nazwy wskaźników	20
3.2	Implementacja interfejsu użytkownika	20
3.2.1	Utworzenie projektu Django	20
3.2.2	Konfiguracja aplikacji	20
3.2.3	Struktura projektu	20
	Główny katalog projektu	21
3.2.4	Aplikacja Parsera	21
3.2.5	Aplikacja Django	21
3.2.6	Rozwiązania zastosowane w projekcie, biblioteki	22
4	Dokumentacja użytkownika	23
4.1	Uruchamianie aplikacji	23
4.2	Strona główna	23
4.3	Dokumenty	25
4.4	Dokumenty ALAB	26

4.5	Wykresy	27
4.6	Lista pacjentów	27

Rozdział 1

Wstęp

1.1 Cel i zakres pracy

Celem pracy było zaprojektowanie systemu przechowującego dokumenty pacjentów i umożliwiającego lekarzowi wgląd we wszystkie dokumenty przechowywane w bazie dla danego pacjenta, w szczególności w dokumenty badania ALAB związane z badaniami krwi i moczu. Lekarz po zalogowaniu do systemu byłby w stanie wgrać dokument PDF i na jego podstawie wygenerować odpowiedni wpis w bazie danych dzięki któremu można by porównywać wyniki pacjenta z badań w różnych okresach czasu. Każdy pacjent w bazie miałby swój własny zbiór dokumentów, i dzięki temu możliwe byłoby szybkie dostanie się przez użytkownika aplikacji do konkretnych wyników danego pacjenta.

Projekt został wykonany przy użyciu języka Python i frameworka Django. Jako bazę danych wykorzystano PostgreSQL. Zespół, który podjął się wykonania projektu składał się z czterech osób: Jakub Frieske, Bartosz Kowalski, Wojciech Marciniak i Bartosz Paulewicz.

Aplikacja umożliwia wgrywanie i przechowywanie wszelkiego rodzaju dokumentów medycznych, pozwalając przede wszystkim w szybki sposób na ich wgląd do nich. W przypadku dokumentów w formacie badania ALAB aplikacja umożliwia jeszcze dodatkowe czynności, takie jak przedstawienie wszystkich wyników zawartych w takich dokumentach w postaci tabelarycznej oraz przedstawianie tych wyników na wykresach, których widok może być dostosowany do potrzeb użytkownika. Podczas wgrywania takiego dokumentu badania ALAB jego zawartość jest parsowana, a następnie użytkownik zatwierdza poszczególne wyniki. W razie potrzeby istnieje oczywiście możliwość korekty danego wyniku lub edycji danych, i po wykonaniu tych czynności można zatwierdzić zapis wczytywanego dokumentu do bazy. Po zgromadzeniu kilku dokumentów danego pacjenta aplikacja umożliwia dzięki wykresom obserwowanie zmian w czasie wybranych przez użytkownika parametrów, dzięki czemu analiza wartości gromadzonych w trakcie badań może być przeprowadzona sprawnie i w bardzo prosty sposób. Formatów różnego rodzaju badań w medycynie jest oczywiście bardzo dużo: jednak aplikacja przeznaczona jest głównie do obsługi formatu badań ALAB, który to format jest sprawozdaniem z badań laboratoryjnych dotyczących m.in. krwi i moczu. To dane z tych dokumentów są parsowane przez aplikację i umożliwiające jest ich porównywanie oraz wizualizacja.

Rozdział 2

Analiza wymagań

W poniższym rozdziale poruszona zostanie szczegółowa analiza wymagań projektu. Wymienione zostaną wszystkie aspekty, które należało wziąć pod uwagę podczas analizy problemu, wykonania projektu systemu oraz jego implementacji. Przedstawiono również wygląd i podział dokumentu w formie wyprodukowanym przez badanie ALAB.

2.1 Wymagania funkcjonalne

Aplikacja opisywana w niniejszej dokumentacji jest przeznaczona dla użytkowników będących lekarzami lub innymi pracownikami ośrodka wykonującego lub analizującego badania krwi i moczu ALAB, którzy chcieliby w łatwy sposób przechowywać dokumenty pacjentów i mieć do nich wgląd. Poza tym jedną z najważniejszych funkcji byłaby możliwość przekształcenia wgranego dokumentu PDF wygenerowanego przez ALAB do postaci tabelarycznej w programie. Możliwy byłby również wgląd do tych danych nie tylko w postaci tabeli ale również wykresów, porównujących dane z kilku różnych badań. Dostępna byłaby opcja wyboru jednego lub kilku wskaźników do porównania.

Poniżej wypunktowany został zbiór wymagań funkcjonalnych oraz struktury danych w aplikacji:

- Pacjent – aplikacja powinna pozwalać użytkownikowi (czyli lekarzowi) na tworzenie w systemie pacjentów, dla których będzie możliwe dodawanie dokumentów medycznych oraz możliwość wglądu do umieszczonych w tabelach oraz zobrazowanych na wykresach danych dotyczących badań ALAB. Pacjenci rozróżniani są przez numery PESEL i posiadają podstawowe dane jak imię, nazwisko, data urodzenia i płeć, po których można ich wyszukać.
- Dokument PDF – aplikacja powinna umożliwiać wgrywanie wszelkiego rodzaju dokumentów medycznych związanych z danym pacjentem, w różnych formatach, takich jak dokumenty, zdjęcia itd. Szczególnym przypadkiem są dokumenty PDF, które można podejrzeć bezpośrednio w aplikacji bez dodatkowego pobierania takich dokumentów przez użytkownika.
- Auto-parsowanie – po wgraniu dokumentu dotyczącego badania ALAB dane zawarte w takim dokumencie są parsowane i zatwierdzane przez użytkownika, po opcjonalnej korekcie manualnej. Korekty można dokonać do pojedynczych wskaźników sparsowanych danych, np. w wypadku wykrycia błędnej wartości lub w przypadku chęci poprawienia jednostki danej wartości na inną.
- Składowanie wyniku – każdy wynik parsowania jest zapisywany i przechowywany w bazie danych. Wyniki z różnych badań w obrębie jednego dokumentu są przyporządkowywane do

odpowiedniego rodzaju badania. Każdy z wyników posiada nazwę, wartość zmierzoną, jednostkę i ewentualnie dodatkowe dane związane z tym wynikiem takie jak zakres referencyjny.

- Rekonstrukcja wyniku – aplikacja umożliwi wgląd we wszystkie zapisane w bazie danych wyniki. Wyniki te są łatwo dostępne z widoku zbioru badań dotyczących danego pacjenta.
- Wykresy w czasie – wszystkie wyniki dotyczące badań ALAB poszczególnych pacjentów mogą zostać zobrazowane w czasie, wizualizacja może dotyczyć pojedynczej wartości lub kilku naraz, w oknie czasowym definiowalnym przez użytkownika.
- Przechowywanie dokumentów innych niż PDF – aplikacja powinna umożliwić wgrywanie różnego rodzaju zdjęć, dokumentów oraz różnych plików innego typu, które również zawierają dane medyczne.

2.2 Wymagania pozafunkcjonalne

Poniżej wypisane zostały wymagania pozafunkcjonalne w projekcie, związane z technologiami, które miały zostać wykorzystane do implementacji aplikacji:

- Język programowania Python - jeden z najpopularniejszych języków programowania na świecie
- Interfejs Django - popularny framework do tworzenia interfejsu użytkownika
- Baza danych PostgreSQL - jeden z najlepszych na rynku otwartych systemów do zarządzania relacyjnymi bazami danych
- Docker - narzędzie pozwalające na umieszczenie aplikacji w wirtualnym kontenerze dzięki czemu zapewniona jest jej przenośność

2.3 Opis dokumentu badania ALAB

W niniejszej sekcji opisano strukturę dokumentu badania ALAB, który jest pewnego rodzaju standardem przedstawiania tabelarycznego wyników badań krwi i moczu.

Dokument badania ALAB, przedstawiony na rys. 2.1 jest zbiorem informacji i wyników dotyczących konkretnego badania lub kilku badań przeprowadzonych po kolei u danego pacjenta. Poniżej zostaną omówione poszczególne sekcje dokumentu, zawartość, którą każda z tych sekcji może zawierać, oraz zaproponowany przez twórców aplikacji podział tego dokumentu, który odzwierciedla projekt bazy danych.

2.3.1 Zawartość dokumentu ALAB

Dokument badania ALAB jest podzielony na sekcje. Nagłówek dokumentu może zawierać różne elementy dotyczące laboratorium, w którym wykonane zostało badanie, takie jak logo laboratorium, kod kreskowy dokumentu oraz zbiór informacji takich jak adres laboratorium oraz ogólne dane kontaktowe z pracownią. Po nagłówku następuje sekcja z danymi pacjenta oraz danymi dotyczącymi badania lub zbioru badań opisywanych przez dokument, zatytułowana „Sprawozdanie z badań laboratoryjnych”. Kolejne sekcje związane są z poszczególnymi typami badań, takimi jak: Hematologia, Biochemia, Immunochemia oraz Analityka moczu. Dokument kończy zestawienie zespołu lekarzy wykonujących badania oraz osób zatwierdzających, oraz informacje dotyczące przeznaczenia dokumentu.



ARZKACZ



ALAB Laboratoria Sp.z o.o.
00-739 Warszawa, ul. Stępińska 22/30
Laboratorium Analiz Lekarskich ALAB Poznań Starolecka
61-360 Poznań, ul. Starolecka 42
tel. 61 871-54-94
e-mail: cs.poznan@alab.com.pl
www.alabalaboratoria.pl



AM 018

Sprawozdanie z badań laboratoryjnych

Psiurbot Alfred

PESEL: 90012409752
Data urodzenia: 24-01-1990
Ident. pacjenta: PESEL_90012409752
Płeć: męczyzna
Ident. dokumentu zlecenia: 99920210991317203177
Nr/data w księdze prac. diagn.: 550 / 26-01-2022

Zleceniodawca: Prakt. Lek. Rodz. M.Kaczmarek i K.Kaczmarek spółka partnerska, Poznań
Miejsce odesłania wyniku: Zleceniodawca
Lekarz zlecający: lekarz Katarzyna Kaczmarek, numer: 1760615
Data i godz. rejestracji zlecenia: 26-01-2022 12:29

Data wykonania badania: 26-01-2022

Hematologia			
Nazwa badania	Wynik badania	Zakres referencyjny	Dokumenty odniesienia
Morfologia krwi (C55)			
Material: Krew żylna pełna (EDTA), data i godz. pobrania: nie podano, data i godz. przyjęcia: 26-01-2022 12:44			
Leukocyty (WBC)	4,14 10 ⁹ /L	4,0 — 10,0	
Eryocyty (RBC)	4,52 10 ¹² /L	4,50 — 5,50	Instrukcja Horiba - 07-2018
Hemoglobina (HGB)	14,8 g/dL	14,0 — 18,0	
Hematokryt (HCT)	42,9 %	40,0 — 54,0	
<u>Wskaźniki czerwonych krwinek:</u>			
Średnia objętość erycyty (MCV)	94,9 fL	80,0 — 100,0	
Średnia masa HGB w erytrocytach (MCH)	32,7 pg	27,0 — 34,0	
Średnie stężenie HGB w erytrocytach (MCHC)	34,4 g/dL	31,0 — 37,0	
Wskaźnik anizocytozy erytrocytów (RDW)	13,2 %	11,0 — 16,0	
Płytki krwi (PLT)	↓ 134 10 ⁹ /L	150,0 — 450,0	
<u>Wskaźniki płytkowe:</u>			
Średnia objętość płytek krwi (MPV)	10,1 fL	6,0 — 11,0	
Rozdział automatyczny leukocytów:			
Neutrocyty (NEU%)	↓ 41,9 %	57,0 — 65,0	
Limfocyty (LYMPH%)	↑ 49,0 %	20,0 — 45,0	
Monocyty (MON%)	7,0 %	3,0 — 8,0	
Eozynocyty (EOS%)	1,7 %	1,0 — 5,0	
Bazocyty (BASO%)	↓ 0,4 %	0,5 — 1,6	
Neutrocyty (NEU)	↓ 1,74 10 ⁹ /L	1,80 — 7,00	
Limfocyty (LYMPH)	2,01 10 ⁹ /L	0,80 — 4,00	
Monocyty (MON)	0,29 10 ⁹ /L	0,20 — 0,80	
Eozynocyty (EOS)	0,07 10 ⁹ /L	0,04 — 0,40	

2.3.2 Opis głównych sekcji

Nagłówek

W nagłówku znajdują się najważniejsze dane: identyfikujące pacjenta oraz badanie. Rysunek 2.2 przedstawia jak może wyglądać ta sekcja.

Sprawozdanie z badań laboratoryjnych			
Psiurbot Alfred			
PESEL:	90012409752	Zleceńodawca:	Prakt. Lek. Rodz. M.Kaczmarek i K.Kaczmarek spółka partnerska, Poznań
Data urodzenia:	24-01-1990	Miejsce odesłania wyniku:	Zleceńodawca
Ident. pacjenta:	PESEL_90012409752	Lekarz zlecający:	lekarz Katarzyna Kaczmarek, numer: 1760615
Płeć: mężczyzna	mężczyzna	Data i godz. rejestracji zlecenia:	Data i godz. rejestracji zlecenia: 26-01-2022 12:29
Ident. dokumentu zlecenia:	99920210991317203177	Data wykonania badania:	26-01-2022
Nr/data w księdze prac. diagn.:	550 / 26-01-2022		

RYSUNEK 2.2: Przykładowy nagłówek dokumentu badania ALAB

Umieszczone w nagłówku są zawsze najważniejsze dane dotyczące pacjenta:

- PESEL
- Data urodzenia
- Płeć

Na dane związane z badaniem mogą, ale nie muszą składać się: identyfikator dokumentu zlecenia, numer/data w księdze prac. diagn., zleceńodawca, lekarz zlecający, miejsce odesłania wyniku oraz najważniejsza dana czyli **data i godzina badania** (oprócz tego może być jeszcze data i godzina np. rejestracji).

Sam nagłówek może (ale nie musi) powtórzyć się na kolejnych stronach dokumentu, w przypadku gdy jest on wielostronny.

Sekcja dotycząca badania

Każda z sekcji dotyczących badania jest strukturalnie ułożona w następujący sposób: badania są spisane jedno po drugim w następującej strukturze: Nazwa badania z skrótem umieszczonym w nawiasach lub sam skrót. W następnej kolumnie umieszczony jest wynik: wartość liczbową z jednostką opcjonalnie poprzedzoną strzałką wskazującą na to jak interpretować wynik względem zakresu referencyjnego, lub opis słowny wyniku. Kolejną kolumną może być właśnie podany ten zakres referencyjny dla tej wartości jeżeli jest ona liczbą. Ostatnia kolumna jest opcjonalna i może zawierać metodę badania, dane dotyczące wykonania badania, dane lekarza wykonującego badanie, lub inne informacje dotyczące tego konkretnego badania. Każda „krotka” może być również zakończona krótką informacją o metodzie badania lub komentarz dla lekarza w przypadku wartości, które należy w specjalny sposób zinterpretować. Rysunki 2.3, 2.4 i 2.5 przedstawiają kilka przykładowych sekcji.

Sekcja końcowa

Sekcja końcowa może zawierać zestawienie osób wykonujących badania oraz osób zatwierdzających, oraz informacje dotyczące przeznaczenia dokumentu, jej wygląd przedstawia rys. 2.6

Hematologia			
Nazwa badania	Wynik badania	Zakres referencyjny	Dokumenty odniesienia
Materiał: Krew żylna pełna (EDTA), data i godz. pobrania: nie podano, data i godz. przyjęcia: 26-01-2022 12:44			
Morfologia krwi (C55)			
Leukocyty (WBC)	4,14 10 ⁹ /L	4,0 — 10,0	
Erytrocyty (RBC)	4,52 10 ¹² /L	4,50 — 5,50	Instrukcja Horiba - 07-2018
Hemoglobina (HGB)	14,8 g/dL	14,0 — 18,0	
Hematokryt (HCT)	42,9 %	40,0 — 54,0	
Wskaźniki czerwonych krwinek:			
Średnia objętość erytrocyta (MCV)	94,9 fL	80,0 — 100,0	
Średnia masa HGB w erytrocycie (MCH)	32,7 pg	27,0 — 34,0	
Średnie stężenie HGB w erytrocytach (MCHC)	34,4 g/dL	31,0 — 37,0	
Wskaźnik anizocytozy erytrocytów (RDW)	13,2 %	11,0 — 16,0	
Płytki krwi (PLT)	↓ 134 10 ⁹ /L	150,0 — 450,0	
Wskaźniki płytkowe:			
Średnia objętość płytek krwi (MPV)	10,1 fL	6,0 — 11,0	
Rozdział automatyczny leukocytów:			
Neutrocyty (NEU%)	↓ 41,9 %	57,0 — 65,0	
Limfocyty (LYMPH%)	↑ 49,0 %	20,0 — 45,0	
Monocyty (MON%)	7,0 %	3,0 — 8,0	
Eozynocyty (EOS%)	1,7 %	1,0 — 5,0	
Bazocyty (BASO%)	↓ 0,4 %	0,5 — 1,6	
Neutrocyty (NEU)	↓ 1,74 10 ⁹ /L	1,80 — 7,00	
Limfocyty (LYMPH)	2,01 10 ⁹ /L	0,80 — 4,00	
Monocyty (MON)	0,29 10 ⁹ /L	0,20 — 0,80	
Eozynocyty (EOS)	0,07 10 ⁹ /L	0,04 — 0,40	

Środowisko Centrum - © MARCEL S.A.

Data i godzina wydania: 26-01-2022 15:06, strona 1 z 3

Bazocyty (BASO)	↓ 0,02 10 ⁹ /L	0,050 — 0,160	
<i>Badanie wykonano na analizatorze Yumizen H2500 firmy Horiba ABX</i>			

RYSUNEK 2.3: Przykładowa sekcja dotycząca badania - Hematologia

Biochemia			
Nazwa badania	Wynik badania	Zakres referencyjny	Dokumenty odniesienia
Materiał: Krew żylna, surowica, data i godz. pobrania: nie podano, data i godz. przyjęcia: 26-01-2022 12:55			
Sód w surowicy (O35)	142 mmol/L	136 — 145	Instrukcja Abbott - 01-2018
<i>Badania wykonano metodą potencjometrii pośredniej na analizatorze Alinity c firmy Abbott.</i>			
Potas w surowicy (N45)	3,6 mmol/L	3,5 — 5,1	Instrukcja Abbott - 01-2018 Instrukcja Abbott - 02-2018
<i>Badania wykonano metodą potencjometrii pośredniej na analizatorze Alinity c firmy Abbott.</i>			
Wapń całkowity w surowicy (O77)	9,5 mg/dL	8,4 — 10,2 Wapń całkowity w surowicy (mmol/l)	
	2,38 mmol/L	2,10 — 2,55	
<i>Badania wykonano metodą spektrofotometryczną na analizatorze Alinity c firmy Abbott.</i>			
Immunochemia			
Nazwa badania	Wynik badania	Zakres referencyjny	Dokumenty odniesienia
Materiał: Krew żylna, surowica, data i godz. pobrania: nie podano, data i godz. przyjęcia: 26-01-2022 12:55			
Tyreotropina (TSH) trzeciej generacji (L69)	1,313 µIU/mL	0,350 — 4,940	Instrukcja Abbott - 02-2018 A
<i>Badania wykonano metodą chemiluminescencji na analizatorze Alinity i firmy Abbott.</i>			

RYSUNEK 2.4: Przykładowa sekcja dotycząca badania - Biochemia i Immunochemia

Analityka			
Nazwa badania	Wynik badania	Zakres referencyjny	Dokumenty odniesienia
Material: Mocz, data i godz. pobrania: nie podano, data i godz. przyjęcia: 26-01-2022 12:52			
Badanie ogólne moczu (A01)			
Barwa	żółta	żółta	PB402_13_65 wyd.2 z 5.11.2019
Przejrzystość	przejrzysty	przejrzysty	
Odczyn pH	6,5	4,6 — 8,0	
Ciężar właściwy	1,011 g/mL	1,012 — 1,026	
Białko	nieobecne	nieobecne	
Glukoza	nieobecna	nieobecna	
Bilirubina	nieobecna	nieobecna	
Urobilinogen	0,2 mg/dL	0,1 — 1,0	
Ciała ketonowe	nieobecne	nieobecne	
Erytrocyty	nieobecne	nieobecne	
Azotyny	nieobecne	nieobecne	
Leukocyty	nieobecne	nieobecne	
Mikroskopowy osad moczu (A19)			
Nabłonki płaskie	pojedyncze wpw	pojedyncze	
Leukocyty	0 - 2 wpw	0 — 5 wpw	
Erytrocyty	0 - 2 wpw	0 - 3 wpw	
Bakterie	pojedyncze wpw	nieobecne	
Pasma śluzu	pojedyncze wpw	nieobecne	

Badanie wykonano na analizatorze Atellica 1500 firmy Siemens

RYSUNEK 2.5: Przykładowa sekcja dotycząca badania - Analityka

A - badanie cechy i metody akredytowane przez Polskie Centrum Akredytacji; PB - procedura badawcza

Wykonali	Zatwierdzili
a - Głuszek Aleksandra	1 - mgr Aleksandra Głuszek
b - Zbierska Małgorzata	2 - mgr Małgorzata Iwona Zbierska
c - Wójcik Patrycja Ewa	3 - mgr Patrycja Ewa Wójcik

Niniejszy wydruk jest informacją o sprawozdaniu z badania laboratoryjnego. Sprawozdanie zostało sporządzone w postaci elektronicznej i podpisane kwalifikowanym podpisem elektronicznym
Autoryzował: diagnosta laboratoryjny Patrycja Dura nr PWZDL 10221

Bez pisemnej zgody Laboratorium sprawozdanie z badań nie może być powielane inaczej jak tylko w całości. Wyniki badań odnoszą się wyłącznie do próbki, której rodzaj, data i godzina pobrania, data i godzina przyjęcia do badań są identyfikowane w niniejszym sprawozdaniu.

RYSUNEK 2.6: Przykładowa sekcja dotycząca badania - Analityka

Rozdział 3

Opis implementacji

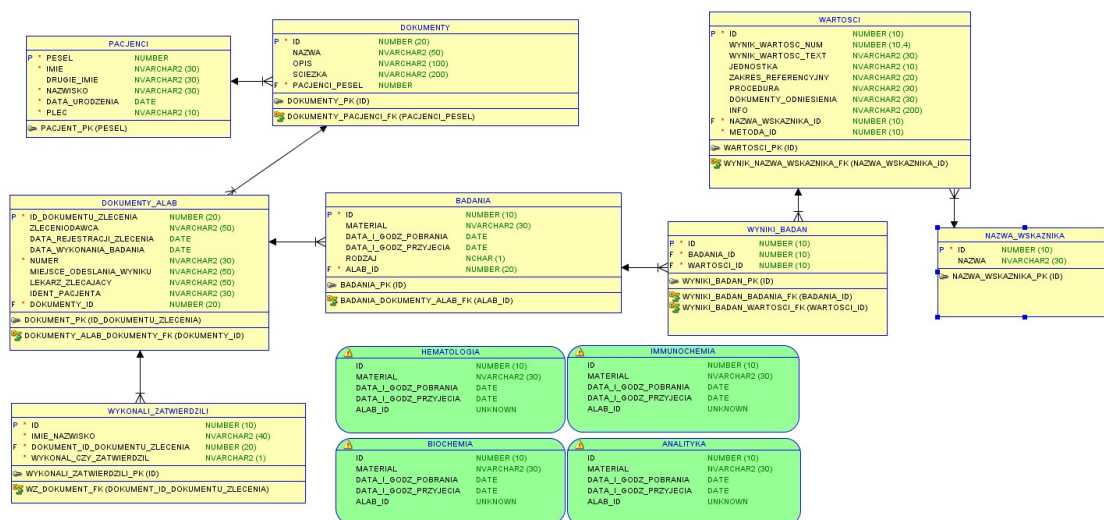
W niniejszym rozdziale opisana została implementacja systemu. Przedstawiono i opisano schemat bazy danych wykorzystywany w projekcie, struktura projektu Django, konfiguracja i krótki opis najważniejszych bibliotek i rozwiązań zastosowanych w projekcie.

3.1 Schemat bazy danych

W poniższej sekcji przedstawiony zostanie zaproponowany przez twórców aplikacji schemat bazy danych oraz dokładny opis każdej z tabel w bazie.

3.1.1 Spis tabel

Rysunek 3.1 przedstawia schemat relacyjnej bazy danych, którą zaprojektowano w ramach systemu.



RYСУNEK 3.1: Przykładowa sekcja dotycząca badania - Analityka

Kluczowe tabele to przede wszystkim tabela Pacjenci: krotka w tej tabeli odpowiada pojedynczemu pacjentowi i zawiera jego najważniejsze dane. Tabela **Dokumenty** odpowiada wszystkim poszczególnym dokumentom pacjentów i innym plikom medycznym (np. zdjęciom), natomiast tabela **Dokumenty_ALAB** jest zawiera dane związane z dokumentem badania ALAB czyli szczególnie traktowanym przez aplikację. Szczegółowy opis każdej z tabel znajduje się w sekcjach poniżej.

3.1.2 Pacjenci

Tabela Pacjenci składa się z następujących pól:

- PESEL – **klucz główny**, różny dla każdego Pacjenta. Jest to 11-cyfrowa liczba, unikalna dla każdej osoby
- Imię – ciąg znaków będący reprezentacją imienia pacjenta.
- *Drugie_Imię* – *opcjonalny* ciąg znaków będący reprezentacją drugiego z imion pacjenta.
- Nazwisko – ciąg znaków będący reprezentacją nazwiska pacjenta.
- Data_urodzenia – data urodzenia pacjenta (z dokładnością do dnia)
- Płeć – pojedynczy znak przyjmujący wartość: „M” lub „K”

3.1.3 Dokumenty

Tabela Dokumenty składa się z następujących pól:

- ID – **klucz główny**, identyfikator dokumentu
- Nazwa – *opcjonalna* nazwa dokumentu (nadana przez użytkownika lub uzupełniona na podstawie nazwy dostarczonego pliku)
- Opis – *opcjonalny* opis dokumentu dodany przez użytkownika (jako ciąg znaków)
- Ścieżka – ścieżka do pliku
- Pacjenci_Pesel – klucz obcy do tabeli Pacjenci, identyfikator pacjenta

3.1.4 Dokumenty ALAB

Tabela Dokumenty_ALAB składa się z następujących pól:

- ID – **klucz główny**, identyfikator dokumentu ALAB
- ID_Dokumentu_zlecenia – numer dokumentu badania ALAB
- Zleceniodawca – podmiot zlecający badanie
- Data_rejestracji_zlecenia – data
- Data_wykonania_badania – data
- Numer – numer w księdze prac. diagn. – ciąg znaków
- Miejsce_odesłania_wyniku – ciąg znaków
- Lekarz_zlecający – ciąg znaków na który może składać się, np. imię i nazwisko oraz jego numer
- Dokumenty_id – klucz obcy do tabeli Dokumenty, identyfikator dokumentu

3.1.5 Wykonali zatwierdzili

Tabela Wykonali_zatwierdzili składa się z następujących pól:

- ID – **klucz główny**, identyfikator osoby wykonującej/zatwierdzającej badanie
- Imie_nazwisko – ciąg znaków zawierający imię i nazwisko
- Dokument_ID_Dokumentu Zlecenia
- Wykonal_czy_zatwierdzil – pojedynczy znak przyjmujący wartość „W” lub „Z”

3.1.6 Badania

Tabela Badania składa się z następujących pól:

- ID – **klucz główny**, identyfikator badania
- Materiał – ciąg znaków, materiał podlegający badaniu
- Data_i_godz pobrania – data
- Data_i_godz przyjecia – data
- Rodzaj – pojedynczy znak przyjmujący symbol wykonywanego badania
- Alab_id – klucz obcy do tabeli Dokumenty_ALAB, identyfikator dokumentu ALAB

Zaproponowano również cztery perspektywy (oznaczone na schemacie kolorem zielonym), które odpowiadają czterem głównym rodzajom badań, których dotyczą te dokumenty: Hematologia, Immunochemia, Biochemia i Analityka, dzięki czemu umożliwiony jest szybki dostęp do konkretnego z badań¹.

3.1.7 Wyniki badań

Tabela Wyniki badań składa się z następujących pól:

- ID – **klucz główny**, identyfikator konkretnego wyniku
- Badania_ID – klucz obcy do tabeli Badania, identyfikator badania
- Wartości_ID – klucz obcy do tabeli Wartości, identyfikator wartości

3.1.8 Wartości

Tabela Wartości składa się z następujących pól:

- ID – **klucz główny**, identyfikator wartości dla wyniku
- Wyniki_wartości_num – wartość numeryczna, używana w przypadku gdy wykryta lub wprowadzona wartość jest liczbą
- Wyniki_wartości_text – wartość tekstowa, używana w przypadku gdy wykryta lub wprowadzona wartość jest tekstem (np. *nieobecne*, *obecne*)
- Jednostka – ciąg znaków
- Zakres_referencyjny – ciąg znaków
- Nazwa_wskaźnika_ID – klucz obcy do tabeli Nazwy wskaźników

¹Przy rozszerzaniu aplikacji można by dodać nowe perspektywy dla innego rodzaju, jeżeli zaszłaby taka potrzeba.

3.1.9 Nazwy wskaźników

Tabela Nazwy wskaźników jest słownikiem do tabeli Wartości:

- ID – **klucz główny**, identyfikator
- Nazwa – ciąg znaków będący nazwą wskaźnika

3.2 Implementacja interfejsu użytkownika

3.2.1 Utworzenie projektu Django

Do implementacji aplikacji wykorzystano framework Django, który można pobrać ze strony <https://www.djangoproject.com/download/>. Podczas rozpoczęcia prac nad implementacją skorzystano z polecenia `python django-admin.py startproject nazwa_projektu`. Polecenie to utworzy katalog `nazwa_projektu`, a w nim cztery pliki:

- `__init__.py`, plik jest pusty, jego istnienie jest informacją dla Pythona, że ma traktować katalog jak pakiet,
- `manage.py`, skrypt do zarządzania utworzonym projektem,
- `settings.py`, skrypt zawierający ustawienia projektu (np. parametry bazy danych),
- `url.py`, skrypt konfiguracyjny zawierający mapowania pomiędzy URL (adres internetowy) a obsługującą go funkcją Pythona.

Do uruchomienia tak utworzonej aplikacji wykorzystano polecenie `python manage.py runserver [numer portu]` (domyślnie może zostać wykorzystany port 8000), po czym wystarczy wejść w przeglądarkę i wpisać w pasek adresu `http://localhost:8000/`.

Samą aplikację utworzono za pomocą polecenia `python manage.py startapp nazwa_aplikacji`. Katalog z aplikacją zawiera:

- `__init__.py`, plik jest pusty, jego istnienie jest informacją dla Pythona, że ma traktować katalog jak pakiet,
- `views.py`, skrypt opisujący wygląd poszczególnych stron,
- `models.py`, skrypt opisujący tabele i relacje w bazie danych,
- `tests.py`, skrypt zawierający testy jednostkowe.

3.2.2 Konfiguracja aplikacji

Konfiguracja aplikacji zawarta została w pliku konfiguracyjnym `.env`, przedstawiony na rys. 3.2. Parametry można zmienić w przypadku chęci skonfigurowania aplikacji do działania w wersji bazy danych online (zakomentowany fragment).

3.2.3 Struktura projektu

W niniejszej sekcji opisana została struktura projektu po skończonej implementacji.


```

.env > ...
1 SECRET_KEY='django-insecure-ua_p6q7we##q3%lod2pj^*-1^w$%)i7-@wx015m(25_%+gcu5g'
2 DEBUG=True
3
4 # Postgres database config
5 DB_NAME='parse_alab'
6 DB_USER='postgres'
7 DB_PASSWORD='postgres'
8 DB_HOST='localhost'
9 DB_PORT='5432'
10
11 # Postgres online database config
12 # DB_NAME='exfcihps'
13 # DB_USER='exfcihps'
14 # DB_PASSWORD='A7QiGfaIF-A9oBp1z_Fkkidwynjge7ak'
15 # DB_HOST='rogue.db.elephantsql.com'
16 # DB_PORT='5432'
17

```

RYSUNEK 3.2: Plik konfiguracyjny

Główny katalog projektu

Główny katalog projektu składa się z następujących komponentów:

- katalog *apps* – w katalogu tym znajduje się aplikacja Django
- katalog *config* – katalog zawierający pliki wymagane do uruchomienia aplikacji, w tym plik *settings.py*, wczytujący ustawienia aplikacji
- katalog *parserALAB* – kod części aplikacji odpowiedzialnej za parsowanie
- katalog *templates* – katalog z szablonami, znajdując się w nim pliki odpowiedzialne za wygląd aplikacji
- plik *manage.py* – plik Django do uruchamiania komend Django
- plik *.env* – plik z konfiguracją aplikacji (aby zmienić konfigurację należy edytować wartości zmiennych zdefiniowanych w tym pliku)
- plik *requirements.txt* – wygenerowany plik z wymaganymi bibliotekami pythona
- plik *Dockerfile* i *docker-compose.yaml* – pliki Dockera

3.2.4 Aplikacja Parsera

Aplikacja parsera jest napisana w języku Python i umieszczona została w folderze *parserALAB* projektu. Plik *Parser.py* zawiera w sobie funkcje pozwalające na parsowanie plików PDF, z czego korzysta później aplikacja Django. Plik *patterns.py* zawiera wzorce jakie mają być rozpoznawane przez parsera. Dodatkowe pliki *extract_metadata.py* i *extract_parameters.py* oraz umieszczony w projekcie przykładowy plik PDF *report.pdf* są plikami testowymi.

3.2.5 Aplikacja Django

Główna aplikacja projektu implementująca wszystkie funkcje wymagane w projekcie została umieszczona w folderze *apps* pod nazwą *patientResult*. Składa się ona z następujących elementów:

- folder *forms* – w folderze tym znajduje się część odpowiedzialna za logikę formularzy
- folder *managers* – w folderze znajdują się pliki z klasami odpowiedzialnym za zarządzanie obiektami modeli Django
- folder *migrations* – w tym miejscu zawarte są migracje bazy danych (w przypadku zmiany struktury modeli w aplikacji to w tym folderze pojawiają się kolejne migracje)
- folder *models* – folder z modelami zdefiniowanymi w aplikacji
- folder *templatetags* – folder z plikiem ze zdefiniowanymi niestandardowymi filtrami i tagami
- folder *views* – folder, w którym znajdują się pliki odpowiadające działaniu aplikacji:
 - *dashresults.py* – rysowanie wykresów
 - *document_alab.py* – wyświetlanie listy dokumentów ALAB, wyświetlanie szczegółów dokumentu ALAB, wczytywanie dokumentu ALAB, generacja danych na podstawie sparsowanego dokumentu i jego zapis
 - *document.py* – wyświetlanie listy dokumentów, dodawanie, usuwanie i edycja dokumentu
 - *errors.py* – błędy w aplikacji
 - *patient.py* – wyświetlanie informacji o pacjencie, wyświetlanie listy pacjenta, dodawanie, usuwanie i edycja pacjentów
 - *shared.py* – logika wyświetlania widoków

3.2.6 Rozwiązania zastosowane w projekcie, biblioteki

- PIPENV – narzędzie do pakietów i tworzenia tzw. *virtual environments*, wykorzystywane do prac nad aplikacją.
- TIKA – parser z oprogramowania Apache Tika <https://tika.apache.org/>, służącego do wykrywania i analizy. W projekcie wykorzystano pythonowy port do Tiki, z dokumentacją na stronie <https://github.com/chris mattmann/tika-python>.
- PLOTLY EXPRESS – biblioteka do rysowania wykresów, szybsza wersja *plotly*. Pozwala w prosty sposób na tworzenie różnego rodzaju interaktywnych wykresów. Link do dokumentacji: <https://plotly.com/python/plotly-express/>.
- DOCKER – kontener do aplikacji.

Rozdział 4

Dokumentacja użytkownika

W rozdziale tym umieszczono szczegółową dokumentację użytkownika. Opisano każdą funkcję systemu, wszystkie dostępne dla użytkownika widoki wraz z opisem udostępnionych opcji. Przedstawione opcje zostały spisane w odpowiednich sekcjach.

4.1 Uruchamianie aplikacji

W celu uruchomienia aplikacji należy mieć zainstalowanego Dockera. Jeżeli Docker nie jest zainstalowany, to można pobrać wersję desktop Windows dostępną pod linkiem:

<https://www.docker.com/products/docker-desktop/>.

Po pobraniu należy uruchomić aplikację, a następnie w terminalu lub wierszu poleceń (np. PowerShell lub zwykły cmd) należy wpisać następujące komendy:

```
docker-compose up
```

```
docker-compose run web python manage.py makemigrations
```

```
docker-compose run web python manage.py migrate
```

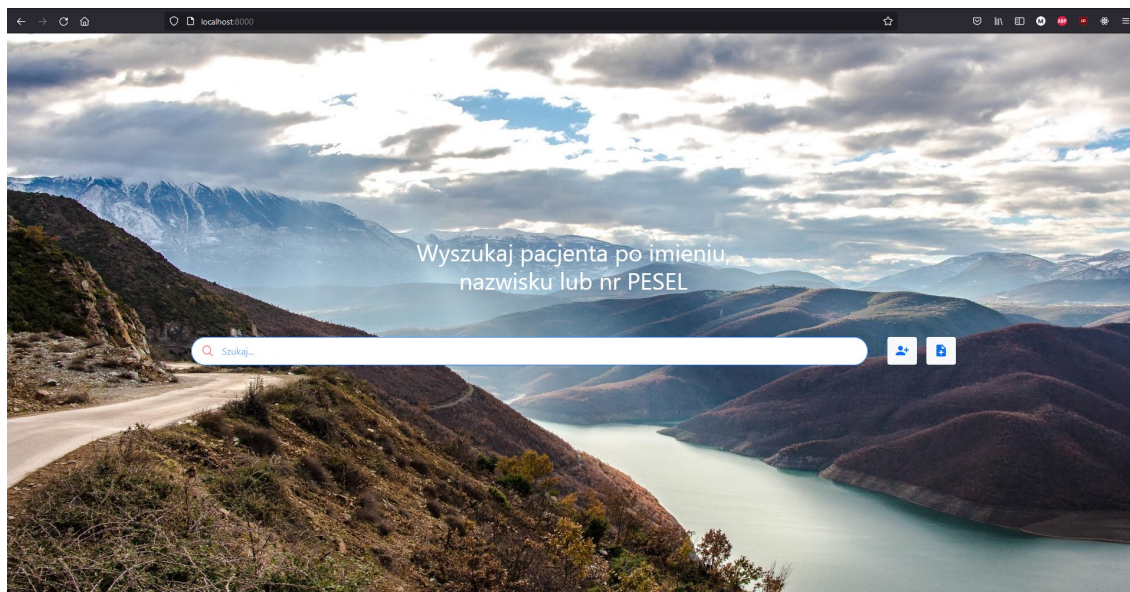
Po wykonaniu pomyślnie powyższych poleceń aplikacja jest gotowa do użycia i jest dostępna pod linkiem `localhost:8000`.

W przypadku, gdy baza już istnieje to nie wykonuje się ponownie migracji i wystarczy jedynie komenda:

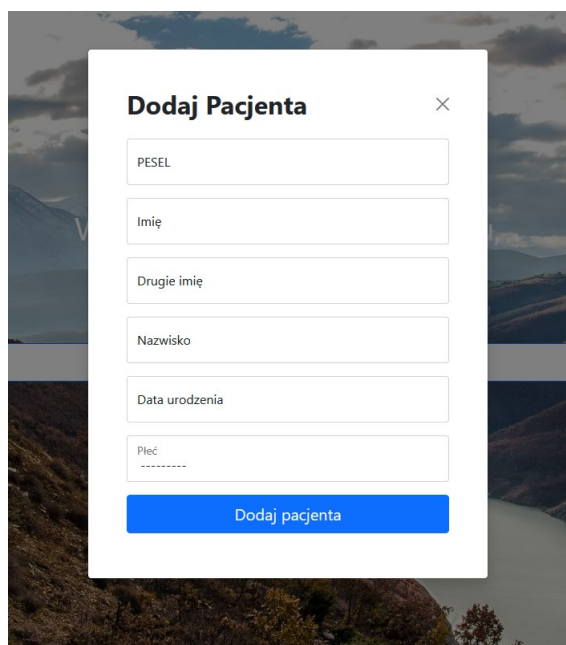
```
docker-compose up
```

4.2 Strona główna

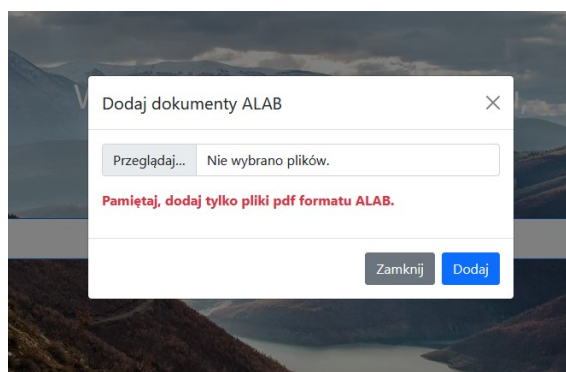
Na stronie głównej, przedstawionej na rys. 4.1 użytkownik ma możliwość wyszukania pacjenta po imieniu, nazwisku albo numerze PESEL. Dostępne są również dwa przyciski: *Dodaj nowego pacjenta* i *Dodaj dokumenty ALAB*. Po kliknięciu pierwszego z nich ukazuje się formularz do wpisywania danych nowego pacjenta (rys. 4.2), natomiast po kliknięciu przycisku do dodania dokumentów ALAB, użytkownik może wybrać pliki, które chce dodać (rys. 4.3).



RYSUNEK 4.1: Dokumenty



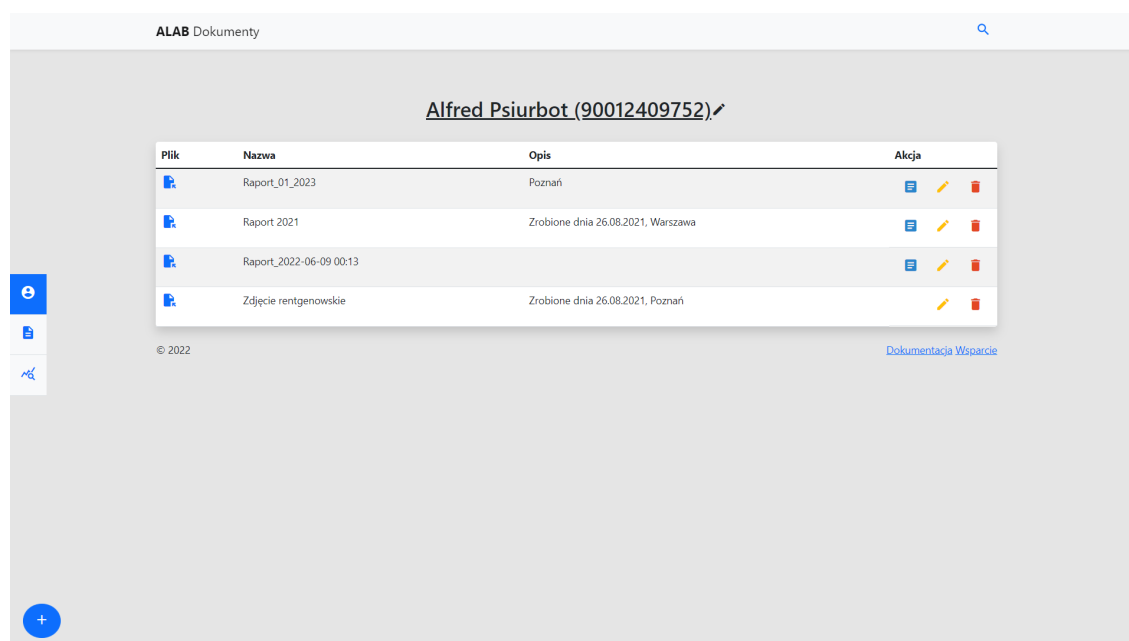
RYSUNEK 4.2: Dokumenty



RYSUNEK 4.3: Dokumenty

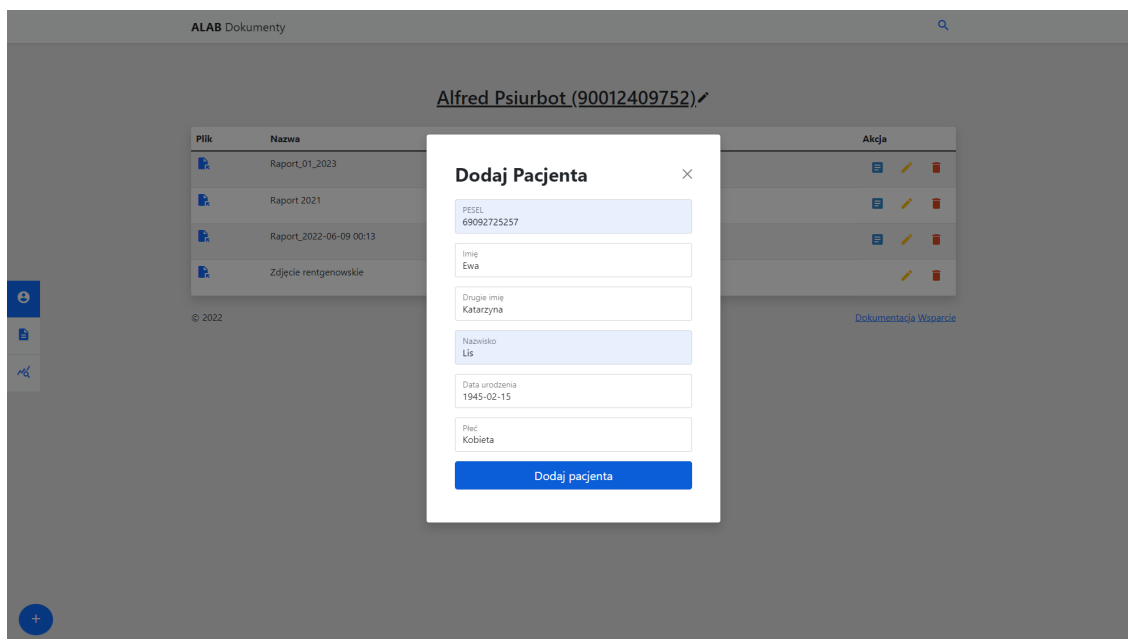
4.3 Dokumenty

Na rysunku 4.4 przedstawiony został widok dokumentów pacjenta. Z poziomu tej strony użytkownik może przede wszystkim dostać się do wszystkich dokumentów, przedstawionych w tabeli. Dla każdej pozycji istnieje możliwość wglądu do niej, edycji oraz jej usunięcia. Poza tym z poziomu tego widoku dostępna jest również edycja danych użytkownika (przycisk edycji po prawej stronie numeru PESEL), co pokazano na rysunku 4.5. Po lewej stronie dostępny jest również panel nawigacyjny pozwalający przejść do 2 innych podstron związanych z pacjentami: dokumentami ALAB oraz wykresami. Przycisk „+” w lewej dolnej części pozwala na dodanie innego pacjenta (rys. 4.6, lub na dodanie dodatkowych dokumentów do bieżącego pacjenta, w tym dokumentów badania ALAB.



RYSUNEK 4.4: Dokumenty

RYSUNEK 4.5: Edycja pacjenta



RYSUNEK 4.6: Dodaj pacjenta

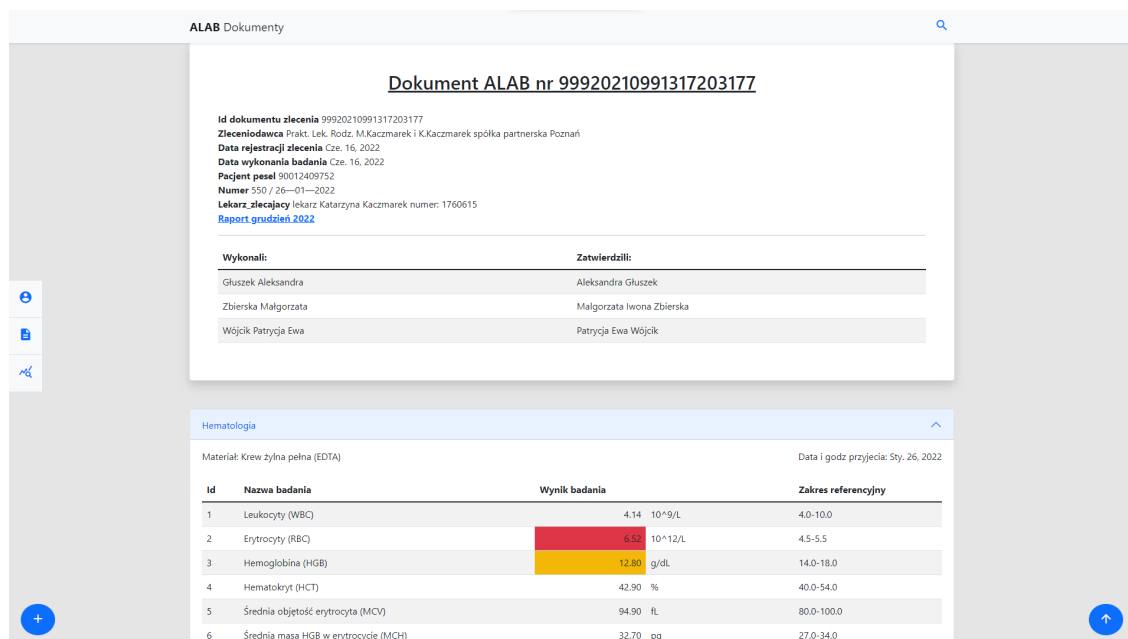
4.4 Dokumenty ALAB

Widok dokumentów ALAB przedstawiony został na rysunku 4.7. Nawigacja jest taka sama jak na stronie *Dokumenty*, ponadto można też kliknąć w dokument aby zobaczyć jego zawartość tabelaryczną, pokazaną na rysunku 4.8.

Id	Zleceniodawca	Data wyk. badania	Numer	Lekarz zlecający
99920210991317203176	Prakt. Lek. Rodz. M.Kaczmarek i K.Kaczmarek spółka partnerska Poznań	2021-06-09	550 / 26—01—2021	lekarz Katarzyna Kaczmarek numer: 1760615
99920210991317203177	Prakt. Lek. Rodz. M.Kaczmarek i K.Kaczmarek spółka partnerska Poznań	2022-06-09	550 / 26—01—2022	lekarz Katarzyna Kaczmarek numer: 1760615
99920210991317203178	Prakt. Lek. Rodz. M.Kaczmarek i K.Kaczmarek spółka partnerska Poznań	2023-06-09	550 / 26—01—2023	lekarz Katarzyna Kaczmarek numer: 1760615

RYSUNEK 4.7: Tabela z wynikami badania ALAB

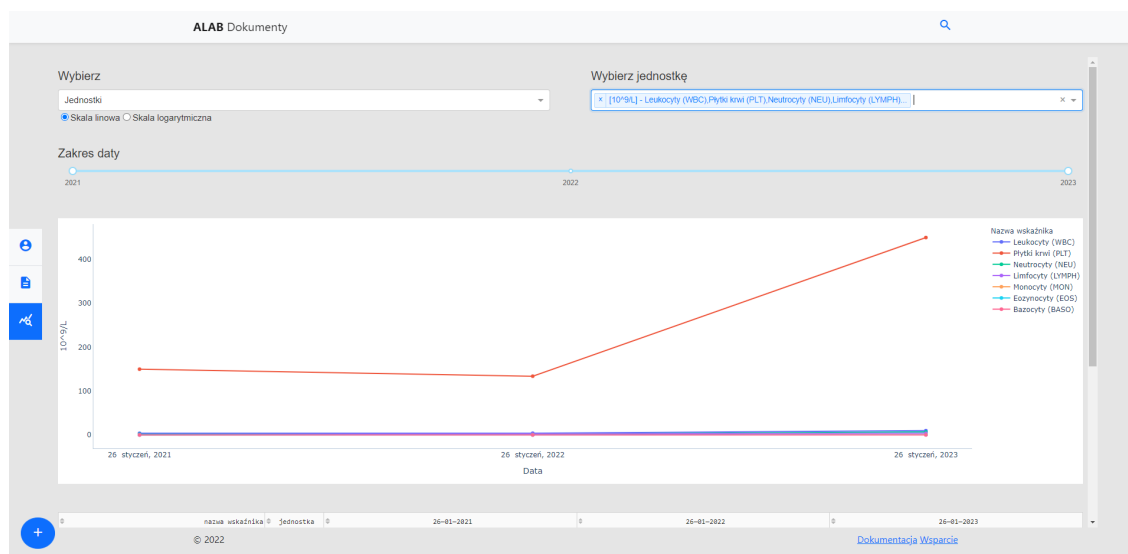
Dodawanie nowych dokumentów ALAB może być automatyczne (wtedy przy dodawaniu należy zaznaczyć opcję „Dodaj automatycznie” lub z możliwością korekty. Jeżeli opcja „Dokument ALAB” zostanie odznaczona, wtedy dokument zostanie potraktowany w zwykły sposób (jako dokument nie-ALAB).



RYSUNEK 4.8: Tabela z wynikami badania ALAB

4.5 Wykresy

Widok wykresów przedstawiony został na rysunku 4.9.



RYSUNEK 4.9: Wykresy

4.6 Lista pacjentów

Widok listy wszystkich pacjentów przedstawia rysunek 4.10.

ALAB Dokumenty

Szukaj

Pesel	Nazwisko	Imiona	Data urodzenia	Płeć
90012409752	Psiurbot	Alfred	1990-01-24	M
71012959822	Lis	Judyta	1971-01-29	K
75092435632	Sobczak	Heronim	1975-09-24	M

© 2022 [Dokumentacja Wsparcie](#)

RYSUNEK 4.10: Widok wszystkich pacjentów



© 2022 Bartosz Paulewicz, Jakub Frieske, Bartosz Kowalski, Wojciech Marciniak

Instytut Informatyki, Wydział Informatyki i Telekomunikacji
Politechnika Poznańska

Skład przy użyciu systemu \LaTeX na platformie Overleaf.