

# Analyse Course 8a PG6

Richard Jansen, Rick Medemblik, Stan Wehkamp, Alex Staritsky

Mei 2017

## Contents

<b>1</b>	<b>Introductie</b>	<b>3</b>
1.1	Aanleiding . . . . .	3
1.2	Doel van het document . . . . .	3
1.3	Doel van de applicatie . . . . .	3
1.4	Doelgroep . . . . .	3
<b>2</b>	<b>Requirements</b>	<b>4</b>
2.1	Doel en focus . . . . .	4
2.2	Functionele eisen . . . . .	4
2.3	Niet-functionele eisen . . . . .	4
<b>3</b>	<b>Use cases</b>	<b>5</b>
3.1	Doel van use cases . . . . .	5
3.2	Actoren . . . . .	5
3.3	Use case 1 - Zoeken met een zoekterm . . . . .	5
3.4	Use case 2 - Visualisatie van de resultaten . . . . .	6
3.5	Use case 3 - Exporteren van de resultaten . . . . .	6
<b>4</b>	<b>Use case diagram</b>	<b>7</b>
4.1	Doel van de use case diagram . . . . .	7
4.2	Actoren . . . . .	7
4.3	Use case diagram . . . . .	7
<b>5</b>	<b>Testplan</b>	<b>8</b>
5.1	Exception handling test . . . . .	8
5.2	Toepassen van testvalues . . . . .	8
5.2.1	Correcte values . . . . .	8
5.2.2	Foutieve values . . . . .	8
5.3	Stresstesting . . . . .	8
5.4	Multiplatform usability . . . . .	9
<b>6</b>	<b>Architectuur</b>	<b>10</b>
6.1	Tier één . . . . .	10
6.2	Tier twee . . . . .	10
6.3	Tier drie . . . . .	10
6.4	Overzicht van het geheel . . . . .	11
<b>7</b>	<b>Client software</b>	<b>12</b>
	<b>Referenties</b>	<b>13</b>

# **1    Introductie**

## **1.1   Aanleiding**

Wetenschappelijke artikelen zoeken en lezen met de hoop om interessante nieuw ontdekkingen te doen kost veel tijd en dat kan efficiënter dacht Ruud Heshof onze opdrachtgever. Hij wil dan ook graag dat er een website komt waarin hij in een oogopslag kan zien welke nieuwe biologische applicaties er zijn voor de zoekterm die hij opgeeft.

## **1.2   Doel van het document**

Het doel van dit document is om de communicatie tussen de bioloog en de informaticus weer te geven. In dit document kan de informaticus lezen welke eisen de bioloog stelt aan de applicatie (de website). Tevens kan de bioloog controleren of alle gewenste onderdelen aanwezig zijn om zijn of haar applicatie (de website) te ontwikkelen.

## **1.3   Doel van de applicatie**

Het doel van de applicatie (de website) is om aan de gebruiker (de bioloog) in een overzichtelijke tabel alle organismes, biologische applicaties en de daarbij horende wetenschappelijke artikelen afkomstig van PubMed behorend bij de zoekterm te tonen.

## **1.4   Doelgroep**

Dit document is geschreven voor biologen. Zij kunnen met behulp van deze applicatie (de website) een zoekterm opgeven en dan in een overzichtelijke tabel alle organismes, biologische applicaties en de daarbij horende wetenschappelijke artikelen afkomstig van PubMed zien. Met behulp van deze informatie kunnen zij verder onderzoek gaan doen.

## 2 Requirements

### 2.1 Doel en focus

Het doel van de requirements is om vast te leggen voor zowel bij de werkgever als de projectleider wat de eisen voor de uiteindelijke applicatie zijn. Zo hebben de werkgever en de projectleider altijd een houvast over van wat afgesproken is, mocht er een discussie ontstaan.

De focus bij dit project ligt bij het maken van een makkelijk te gebruiken webapplicatie die de resultaten overzichtelijk weergeeft.

### 2.2 Functionele eisen

Table 1: Functionele eisen

ID	Use case koppeling	Beschrijving	Prioriteit	Moscow
F1	Zoeken met een zoekterm	De gebruiker kan door activatie van een button het textminen starten.	Hoog	M
F2	Visualisatie van resultaten	De gebruiker kan in een drop down menu Een van de gevonden lipiden selecteren Waarna een tabel verschijnt met dat peptide Het organisme, applicatie en een link naar de bron. ORF's van een gegeven DNA sequentie voorspelt.	Hoog	M
F3	Exporteren van resultaten	De gebruiker heeft de mogelijkheid om de Gegevens op te slaan in een database.	Hoog	M

### 2.3 Niet-functionele eisen

Table 2: Niet functionele eisen.

ID	Beschrijving	Prioriteit	Moscow
F1	Het programma werkt op Windows	Hoog	M
F2	Het programma is een webapplicatie	Hoog	M
F3	De gebruiker heeft resultaten gevisualiseerd. Het programma is voor een bioloog duidelijk te gebruiken, zonder voorlichting vanuit andere personen	Hoog	M
F4	De webapplicatie is publiek toegankelijk.	Hoog	M

## 3 Use cases

### 3.1 Doel van use cases

Het doel van deze use cases zijn om de gebruiker van de applicatie (de bioloog) en de ontwerper te laten zien op wat voor manieren het uiteindelijke programma gebruikt kan worden. Het gaat dan om een "best practice" van wat een bioloog allemaal moet doen om een bepaald resultaat te verkrijgen. Hierdoor krijgt de bioloog een idee van hoe hij de applicatie kan gebruiken zodra deze gemaakt is. De ontwerper kan dit gebruiken om bepaalde elementen in het programma te zetten zodat de use cases mogelijk zijn voor de bioloog.

### 3.2 Actoren

De actor in deze use cases is de gebruiker. De gebruiker is de persoon die de webapplicatie gaat gebruiken voor onderzoek. De gebruiker moet opdrachten aan het programma kunnen geven om een output te krijgen die hij/zij kan interpreteren.

### 3.3 Use case 1 - Zoeken met een zoekterm

De eerste use case is dat de gebruiker, de bioloog een zoekterm of meerdere zoektermen kan opgeven om het programma te starten.

Table 3: Use case 1: 'Zoeken met een zoekterm'

Naam	Zoeken met een zoekterm
Actor	Gebruiker.
Aanamen	De gebruiker heeft verbinding met het internet.
Beschrijving	(1) Gebruiker voert een zoekterm in het zoekveld van de webpagina. (2) (Optioneel) de gebruiker kiest/verandert één of meerdere parameters voor het programma. (3) De gebruiker drukt op de zoek knop om een aanvraag naar de server te sturen. Op het moment dat de gebruiker op de zoek knop drukt worden de gegevens gecontroleerd. In het geval dat er geen zoekterm ingevuld is treedt de uitzondering [Geen zoekterm ingevuld] op. In het geval dat de parameters niet correct zijn treedt de uitzondering [Onjuiste parameter] op.
Uitzonderingen	[Geen zoekterm ingevuld]. De webpagina geeft een melding dat er geen zoekterm is ingevuld. [Onjuiste parameter]. De webpagina geeft aan dat er iets mis is met de ingevulde parameters.
Resultaat	Een melding van de webpagina dat de zoekopdracht naar de server gestuurd is en de pagina blijft aangeven dat het bezig is met verwerken. Zodra er resultaten zijn geeft de webpagina de resultaten terug.

### 3.4 Use case 2 - Visualisatie van de resultaten

De tweede use case beschrijft hoe de resultaten worden gevisualiseerd in de webapplicatie.

Table 4: Use case: 'Visualisatie van de resultaten'

Naam	Visualisatie.
Actor	Gebruiker.
Aanamen	Er staan resultaten in database.
Beschrijving	1a. De gebruiker kiest om te sorteren op peptiden 2 a. De gebruiker kiest om te sorteren op organisme 3 a. De gebruiker kiest om te sorteren op applicatie. 1,2,3 b. Het programma haalt de resultaten uit de database 1 c Het programma laat een tabel zien met de resultaten gesorteerd op de soorten peptiden 2 c. Het programma laat een tabel zien met de resultaten gesorteerd op de soorten organismen. 3 c. Het programma laat een tabel zien met de resultaten gesorteerd op de soorten applicaties.
Uitzonderingen	1b. Als er geen verbinding gemaakt kan worden met pubmed word er een foutmelding gegeven 1c. Als er geen matches worden gevonden worden gevonden word een notificatie gegeven. 1d. Als er geen verbinding met de database gemaakt kan worden verscheind er een foutmelding.
Resultaten	De resultaten van het textminen zijn opgeslagen in een database.

### 3.5 Use case 3 - Exporteren van de resultaten

De derde use case diagram beschrijft hoe de resultaten van de webpagina geëxporteerd kunnen worden.

Table 5: Use case - Exporteren van de resultaten

Naam	Exporteren van de resultaten
Actor	Gebruiker.
Aanamen	De gebruiker heeft resultaten gevisualiseerd.
Beschrijving	1a De gebruiker klikt op exporteren 1b Een filebrowser opent waarin het bestands type en locate van het bestand geselecteerd kan worden. 1c De gebruiker klikt op opslaan en het bestand word opgeslagen
Uitzonderingen	1c Bij een input output exeption verscheind een foutmelding
Resultaat	Een lokaal opgeslagen bestand.

## 4 Use case diagram

### 4.1 Doel van de use case diagram

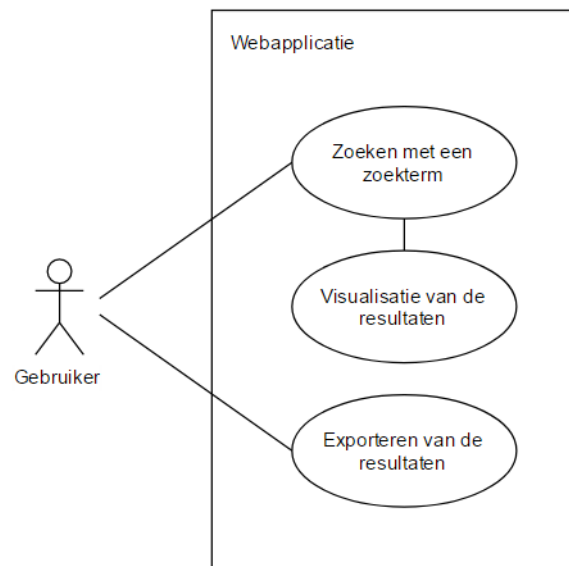
Het doel van de use case diagram is om op een visuele manier alle use cases te weergeven. Hierdoor zijn de use cases wat makkelijker te interpreteren. Het is bedoeld als verheldering van de use cases die beschreven staan in het hoofdstuk 'Use cases'.

### 4.2 Actoren

De actor in deze use case diagram is de gebruiker. De gebruiker is de persoon die de webapplicatie gaat gebruiken voor onderzoek. De gebruiker moet opdrachten aan het programma kunnen geven om een output te krijgen die hij/zij kan interpreteren.

### 4.3 Use case diagram

Figure 1: Use case diagram



## 5 Testplan

Om de werking van het te leveren product te garanderen is een testplan benodigd. Het testplan dient het product bloot te stellen aan een situatie die de situatie van de klant representeert. De onderstaande stappen worden op alle modules toegepast.

### 5.1 Exception handling test

De exception handling aanwezig in de applicatie dient getest te worden op:

- Correct herkennen van fout
- Correct uitvoeren van alternatieve methode na ontdekking van fout
- Juiste hiërarchie van exceptions

Om de bovenstaande punten te testen dient de applicatie blootgesteld te worden aan values die wetenswaardig foutief zijn. In de exception handling wordt er een print gezet die de specifieke fout kan benoemen. Indien dat overeenkomt met de data die toegepast is is de exception handling geslaagd.

### 5.2 Toepassen van testvalues

#### 5.2.1 Correcte values

In overleg met de cliënt wordt er een dataset gemaakt met gebruikersinvoer gerelateerd aan het uiteindelijke gebruik van de applicatie. De testwaarden worden toegepast in de applicatie en indien de uitvoer voldoet aan de eisen wordt de applicatie goedgekeurd op het gebied van werking onder normale omstandigheden.

#### 5.2.2 Foutieve values

Er wordt een dataset aangemaakt met foutieve waarden om de stabiliteit en flexibiliteit van het product te testen. Mits de applicatie geen fatale fout genereert kan ook dit onderdeel goedgekeurd worden.

### 5.3 Stresstesting

Het is mogelijk dat de applicatie meerdere simultane gebruikers kent. Om dit te testen wordt er een applicatie geschreven om het product te testen onder intensieve demand. De responstijd en duur van berekeningen dienen niet significant langer te duren.



## 5.4 Multiplatform usability

Om een zo groot mogelijke groep gebruikers toegang te bieden tot de applicatie is het benodigd om het te publiceren voor alle platforms. Vandaar de keuze voor een web-applicatie. Om dit te testen wordt er gebruik gemaakt van een workstation met virtual machines om het product te testen met meerdere operating systems.

## 6 Architectuur

### 6.1 Tier één

Ook presentation tier genoemd. De presentation tier is verantwoordelijk voor het opvangen van user input en het weergeven van de resultaten. De volgende middelen worden gebruikt voor de presentation tier:

- HTML4[3]
- CSS[4]
- Java Script[4]

Nadat de presentation tier input heeft ontvangen stuurt het de input door naar de logic tier.

### 6.2 Tier twee

Ook logic tier genoemd. De logic tier is verantwoordelijk voor alle berekeningen en het conditioneren van zowel de gebruikersinput als de resultaten. De volgende middelen worden gebruikt door de logic tier:

- Python[7], Flask-package[5] en Biopython(Entrez)[2]
- Bash[1]

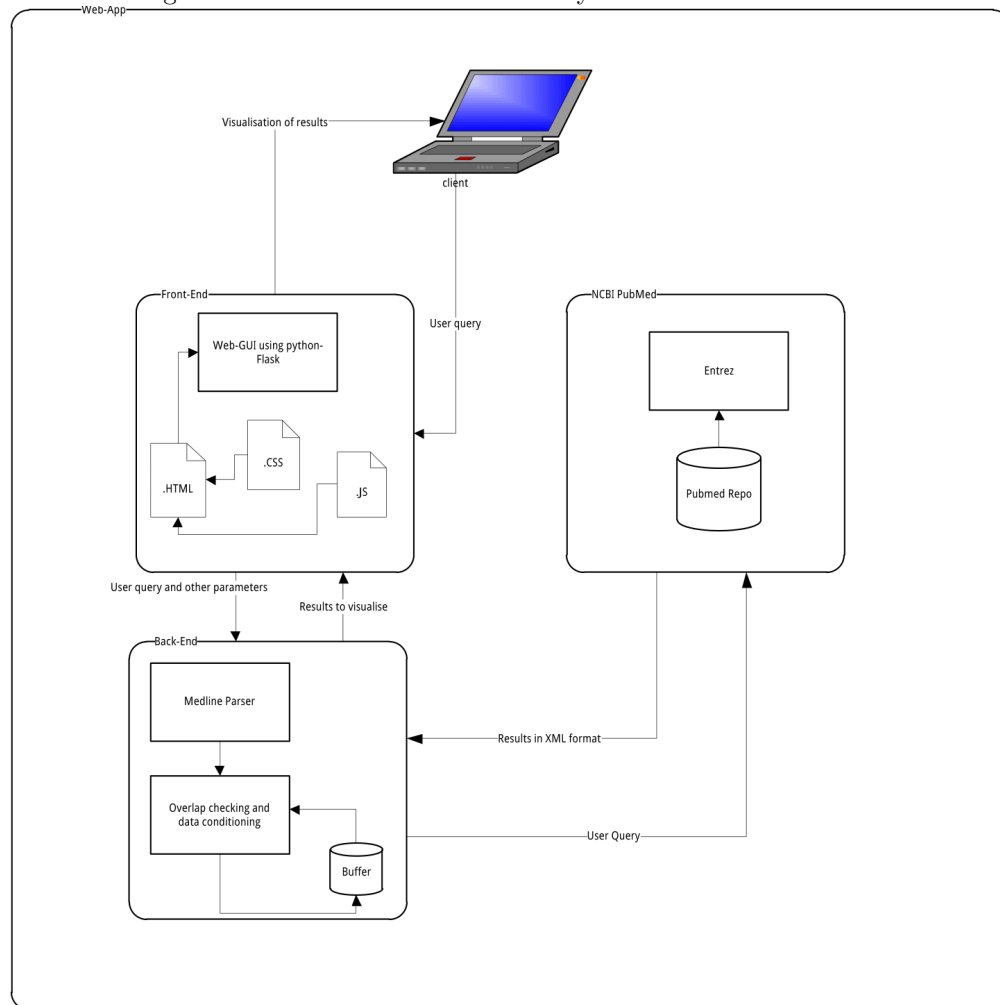
Gebruikersinput wordt verwerkt tot een query voor Entrez[6] en de resultaten uit entrez worden geconditioneerd en naar de presentation tier verstuurd.

### 6.3 Tier drie

Ook data tier genoemd. De data tier is in dit geval de NCBI server met de pubmed-database. De logic tier gebruikt de Entrez API[6] om queries te ontwikkelen die naar de data tier worden gestuurd. De data tier vergaart vervolgens een lijst met resultaten die daarna terug naar de logic tier gaat om verwerkt te worden.

## 6.4 Overzicht van het geheel

Figure 2: Schematisch overzicht van systeemarchitectuur



## 7 Client software

Gezien het feit dat de applicatie web-based is zijn de systeemvereisten aan de cliënt zijde vrij beperkt.

- Chrome/Firefox/Internet Explorer
- Support voor HTML4.4 met CSS en JS

## References

- [1] Ken O Burtch. *Linux shell scripting with bash*. Pearson Higher Education, 2004.
- [2] Peter JA Cock, Tiago Antao, Jeffrey T Chang, Brad A Chapman, Cymon J Cox, Andrew Dalke, Iddo Friedberg, Thomas Hamelryck, Frank Kauff, Bartek Wilczynski, et al. Biopython: freely available python tools for computational molecular biology and bioinformatics. *Bioinformatics*, 25(11):1422–1423, 2009.
- [3] Rick Darnell and Warren Steel. *HTML 4 unleashed*. Macmillan Publishing Co., Inc., 1997.
- [4] Danny Goodman. *Dynamic HTML: The Definitive Reference: A Comprehensive Resource for HTML, CSS, DOM & JavaScript*. " O'Reilly Media, Inc.", 2002.
- [5] Miguel Grinberg. *Flask web development: developing web applications with python*. " O'Reilly Media, Inc.", 2014.
- [6] Donna Maglott, Jim Ostell, Kim D Pruitt, and Tatiana Tatusova. Entrez gene: gene-centered information at ncbi. *Nucleic acids research*, 33(suppl 1):D54–D58, 2005.
- [7] Guido Van Rossum and Fred L Drake Jr. *Python reference manual*. Centrum voor Wiskunde en Informatica Amsterdam, 1995.