

## ONDERZOEKSVOORSTEL

# Het optimaliseren van prestaties voor een Python-applicatie in productie die mogelijke fouten in gegevensbestanden opspoot: een vergelijkende studie en proof-of-concept.

Bachelorproef, 2022-2023

Bjorn Truye

E-mail: [bjorn.truye@student.hogent.be](mailto:bjorn.truye@student.hogent.be)

Co-promotor: S. Lievens (HoGent, [stijn.lievens@hogent.be](mailto:stijn.lievens@hogent.be))

## Samenvatting

In deze bachelorproef zal onderzoek worden gedaan naar de best practices om Python te schrijven, om vervolgens te kunnen toepassen in een bestaand software project van HoGent. Dit project is een Streamlit applicatie geschreven in Python, met een flask-backend en heeft als doel het detecteren van mogelijke fouten in databestanden. De focus van dit onderzoek ligt op het verbeteren van de uitvoersnelheid.

Het onderzoek zal starten met een grondige analyse van de mogelijkheden binnenin Python en waarom Python een al dan niet geschikte taal is voor foutenanalyse van data. Hierbij zal een vergelijkende studie worden gemaakt over de beschikbare tools binnen Python zoals logging, testing, distributie van de applicatie, exception handling, scalability ... Vervolgens zal onderzoek worden gedaan naar de beschikbare externe middelen en andere alternatieven buiten de Python taal zoals open-source frameworks en node packages. De focus hierbij ligt terug op het verbeteren van de uitvoersnelheid om data te analyseren.

Tenslotte zal bekeken worden of de meest optimale technieken reeds toegepast worden in de Streamlit applicatie van HoGent of niet. Alle bevindingen van het voorgaande onderzoek zullen hierbij worden toegepast.

Dit onderzoek zal niet alleen de performantie optimaliseren voor het project van HoGent, maar kan ook dienen als maatstaf voor andere projecten van dezelfde aard. Dit kan bedrijven helpen in hun keuzeproces als ze een gelijkaardig software project willen opzetten.

De verwachting is dat er ruimte voor verbetering is binnen de HoGent applicatie die voor een snellere uitvoertijd kunnen zorgen.

**Keuzerichting:** Mobile & Enterprise development

**Sleutelwoorden:** Python, World Wide Web,  $\lambda$ -calculus

## Inhoudsopgave

1	Introductie . . . . .	1
2	State-of-the-art . . . . .	2
3	Methodologie . . . . .	2
4	Verwacht resultaat, conclusie . . . . .	2
	Referenties . . . . .	2

## 1. Introductie

Waarover zal je bachelorproef gaan? Introduceer het thema en zorg dat volgende zaken zeker duidelijk aanwezig zijn:

- kaderen thema
- de doelgroep
- de probleemstelling en (centrale) onderzoeksvraag
- de onderzoeksdoelstelling

Denk er aan: een typische bachelorproef is *toegespast onderzoek*, wat betekent dat je start van-

uit een concrete probleemsituatie in bedrijfscontext, een **casus**. Het is belangrijk om je onderwerp goed af te bakenen: je gaat voor die *ene specifieke probleemsituatie* op zoek naar een goede oplossing, op basis van de huidige kennis in het vakgebied.

De doelgroep moet ook concreet en duidelijk zijn, dus geen algemene of vaag gedefinieerde groepen zoals *bedrijven*, *developers*, *Vlaamingen*, enz. Je richt je in elk geval op it-professionals, een bachelorproef is geen populariserende tekst. Eén specifiek bedrijf (die te maken hebben met een concrete probleemsituatie) is dus beter dan *bedrijven* in het algemeen.

Formuleer duidelijk de onderzoeksvraag! De begeleiders lezen nog steeds te veel voorstellen waarin we geen onderzoeksvraag terugvinden.

Schrijf ook iets over de doelstelling. Wat zie je als het concrete eindresultaat van je onderzoek, naast de uitgeschreven scriptie? Is het een proof-of-concept, een rapport met aanbevelingen, ... Met welk eindresultaat kan je je bachelorproef als een succes beschouwen?

## 2. State-of-the-art

Hier beschrijf je de *state-of-the-art* rondom je gekozen onderzoeksdomein, d.w.z. een inleidende, doorlopende tekst over het onderzoeksdomein van je bachelorproef. Je steunt daarbij heel sterk op de professionele *vakliteratuur*, en niet zozeer op populariserende teksten voor een breed publiek. Wat is de huidige stand van zaken in dit domein, en wat zijn nog eventuele open vragen (die misschien de aanleiding waren tot je onderzoeksvraag!)?

Je mag de titel van deze sectie ook aanpassen (literatuurstudie, stand van zaken, enz.). Zijn er al gelijkaardige onderzoeken gevoerd? Wat concluderen ze? Wat is het verschil met jouw onderzoek?

Verwijs bij elke introductie van een term of bewering over het domein naar de vakliteratuur, bijvoorbeeld (Hykes, 2013)! Denk zeker goed na welke werken je refereert en waarom.

Draag zorg voor correcte literatuurverwijzingen! Een bronvermelding hoort thuis *binnen* de zin waar je je op die bron baseert, dus niet er buiten! Maak meteen een verwijzing als je gebruik maakt van een bron. Doe dit dus *niet* aan het einde van een lange paragraaf. Baseer nooit teveel aansluitende tekst op eenzelfde bron.

Als je informatie over bronnen verzamelt in JabRef, zorg er dan voor dat alle nodige info aanwezig is om de bron terug te vinden (zoals uitvoerig besproken in de lessen Research Methods).

Je mag deze sectie nog verder onderverdelen in subsecties als dit de structuur van de tekst kan verduidelijken.

## 3. Methodologie

Hier beschrijf je hoe je van plan bent het onderzoek te voeren. Welke onderzoekstechniek ga je toepassen om elk van je onderzoeksvragen te beantwoorden? Gebruik je hiervoor literatuurstudie, interviews met belanghebbenden (bv. voor requirements-analyse), experimenten, simulaties, vergelijkende studie, risico-analyse, PoC, ...?

Valt je onderwerp onder één van de typische soorten bachelorproeven die besproken zijn in de lessen Research Methods (bv. vergelijkende studie of risico-analyse)? Zorg er dan ook voor dat we duidelijk de verschillende stappen terug vinden die we verwachten in dit soort onderzoek!

Vermijd onderzoekstechnieken die geen objectieve, meetbare resultaten kunnen opleveren. Enquêtes, bijvoorbeeld, zijn voor een bachelorproef informatica meestal **niet geschikt**. De antwoorden zijn eerder meningen dan feiten en in de praktijk blijkt het ook bijzonder moeilijk om voldoende respondenten te vinden. Studenten die een enquête willen voeren, hebben meestal ook geen goede definitie van de populatie, waardoor ook niet kan aangetoond worden dat eventuele

resultaten representatief zijn.

Uit dit onderdeel moet duidelijk naar voor komen dat je bachelorproef ook technisch voldoende diepgang zal bevatten. Het zou niet kloppen als een bachelorproef informatica ook door bv. een student marketing zou kunnen uitgevoerd worden.

Je beschrijft ook al welke tools (hardware, software, diensten, ...) je denkt hiervoor te gebruiken of te ontwikkelen.

Probeer ook een tijdschatting te maken. Hoe lang zal je met elke fase van je onderzoek bezig zijn en wat zijn de concrete *deliverables* in elke fase?

## 4. Verwacht resultaat, conclusie

Hier beschrijf je welke resultaten je verwacht. Als je metingen en simulaties uitvoert, kan je hier al mock-ups maken van de grafieken samen met de verwachte conclusies. Benoem zeker al je assen en de onderdelen van de grafiek die je gaat gebruiken. Dit zorgt ervoor dat je concreet weet welk soort data je moet verzamelen en hoe je die moet meten.

Wat heeft de doelgroep van je onderzoek aan het resultaat? Op welke manier zorgt jouw bachelorproef voor een meerwaarde?

Hier beschrijf je wat je verwacht uit je onderzoek, met de motivatie waarom. Het is **niet** erg indien uit je onderzoek andere resultaten en conclusies vloeien dan dat je hier beschrijft: het is dan juist interessant om te onderzoeken waarom jouw hypothesen niet overeenkomen met de resultaten.

## Referenties

Hykes, S. (2013, maart 21). *The future of Linux Containers (PyCon 2013)*. Verkregen 1 september 2016, van <https://www.youtube.com/watch?v=wW9CAH9nSLs>