

ONDERZOEKSVOORSTEL

POC: Linter voor BibLaTeX.

Bachelorproef, 2023-2024

Tristan Cuvelier

E-mail: tristan.cuvelier@student.hogent.be

Co-promotor: B. Van Vreckem (HoGent, bert.vanvreckem@hogent.be)

Samenvatting

In deze bachelorproef wordt een BibLaTeX linter ontwikkeld, specifiek gericht op het verbeteren van bronvermelding in LaTeX-documenten voor studenten en academici aan HoGent. Het project pakt het wijdverspreide probleem van onjuiste bronvermeldingen aan en vervult een bestaande behoefte, aangezien huidige linters zoals die van Pez Cuckow niet voldoen. De proef omvat een methodologie bestaande uit literatuurstudie, technische analyse en de ontwikkeling van een prototype gedurende een periode van 12 weken. Het beoogde resultaat is een open-source BibLaTeX linter prototype dat niet alleen de precisie van bronvermeldingen verhoogt, maar ook bijdraagt aan de algemene kwaliteit van academisch werk, zowel binnen HoGent als daarbuiten.

Keuzerichting: Mobile & Enterprise development

Sleutelwoorden: Proof of Concept, Linter, BibLaTeX

Inhoudsopgave

1	Introductie	1
2	State-of-the-art	2
3	Methodologie	2
3.1	Fase 1: Literatuurstudie	2
3.2	Fase 2: Technische Analyse en Experimenten	2
3.3	Fase 3: Software Ontwikkeling (PoC - Proof of Concept)	2
3.4	Algemeen + Fase 4	3
4	Verwacht resultaat, conclusie	3
	Referenties	3

1. Introductie

LaTeX (L^AT_EX) uitspraak: la-tech; is een populair softwaresysteem in de wetenschappelijke wereld omdat het uitblinkt in het zetten van technische documenten, en beschikbaar is voor bijna alle computersystemen (Oetiker e.a., 2023).

Niet alleen in de wereld van de wetenschap wordt LaTeX gebruikt. Ook studenten op hogescholen en universiteiten maken er gebruik van voor het schrijven van bachelor- en/of masterproeven.

Bij het schrijven van, al dan niet wetenschappelijke, teksten is het uitermate belangrijk om aan een correcte vorm van bronvermelding te doen.

Binnen LaTeX zijn er verschillende manieren om dit aan te pakken. Eén van deze manieren is met behulp van BibLaTeX, een package speciaal gebouwd voor deze taak.

Studenten te HoGent dienen gebruik te maken van deze combinatie bij het schrijven van hun bachelorproef. Ondanks dat lectoren veel moeite steken in het bondig toelichten van het correcte

gebruik, worden er nog veel fouten gemaakt op het correct bijhouden van bronnen. Op deze groep zal deze bachelorproef zich focussen, met behulp van een statische analysetool, ook wel linter genoemd, zouden er al veel van de herhalende fouten voorkomen kunnen worden. Een linter is een programma dat broncode of gestructureerde dataformaten kan controleren op stijl, syntax en logische fouten (Kamunya, 2023).

Het zou dus uitermate geschikt zijn om de studenten een linter te laten gebruiken om hen zo te helpen bij het voorkomen of opsporen van de gemaakte fouten. Zo dienen lectoren hen niet keer op keer op dezelfde fouten te wijzen.

BibLaTeX is voortkomende uit BibTeX en biedt meer opties om bibliografieën en citaten te configureren (Cassidy, 2013). Hoewel er voor BibTeX reeds een linter bestaat, is deze niet compatibel met BibLaTeX.

Het doel van deze bachelorproef is om een proof of concept, analyse & de software-architectuur uit te werken voor een BibLaTeX linter en er een prototype voor te schrijven in een passende programmeertaal.

Concreet betekent dit:

- De lijst van gewenste functionele en niet-functionele requirements aanvullen en structureren naar prioriteiten
- De werking van bestaande linters bestuderen als inspiratiebron
- Een gemotiveerde keuze maken voor de te gebruiken programmeertaal en eventuele libraries
- Een prototype met een minimale set van linting-regels implementeren

- Unit tests schrijven met zo compleet mogelijke code coverage
- CI-pipeline opzetten voor packaging en testing
- Documentatie schrijven voor het gebruik en uitbreiden van de linting-regels

Als eindresultaat voor deze bachelorproef zal er een open-source prototype opgesteld worden waaraan vrijwilligers verder kunnen werken.

2. State-of-the-art

Op het ogenblik van het schrijven van dit bachelorproefvoorstel, zijn er nog geen *optimale* BibLaTeX-Linters beschikbaar. De enige beschikbare linter die bestaat voor BibLaTeX voor dit moment, staat op een GitHub-repository van Pez Cuckow. Deze zou wel werkende zijn, maar lijkt niet zo optimaal op het eerste zicht. Er is dus duidelijk mogelijkheid tot verbetering. De BibLaTeX-Linter van Pez Cuckow is geschreven in Python en heeft een webinterface (Cuckow, 2022). Tijdens het onderzoek naar de werking hiervan, werd er niet in geslaagd om deze werkende te krijgen. Dit is eventueel een opdracht voor tijdens de effectieve uitvoer. Hierdoor mag ook al de conclusie getrokken worden dat de bestaande Linter voor BibLaTeX dus nog zeker niet optimaal is op vlak van *gebruiksvriendelijkheid*, gezien de nodige tijd om hem werkende te krijgen.

3. Methodologie

Dit onderzoek zal opgesplitst worden in meerdere fasen. Voor het uitvoeren van dit onderzoek zal wellicht een computer het enige zijn dat nodig is, met uitzondering op technisch inzicht en logisch redeneren. Het eindresultaat voor dit onderzoek zal een open-source prototype zijn van een BibLaTeX-linter waaraan vrijwilligers verder kunnen werken.

Voor het onderzoek zouden 14 weken ter beschikking gesteld worden. Het is de bedoeling om de fasen te verdelen over deze weken en om alsnog een week of twee over te houden ter reserve voor moesten er fouten bij de schattingen ingeslopen zijn. De inschatting van de benodigde tijd zal bepaald worden op basis van het aantal onderzoek dat er moet gebeuren en van het verwachte resultaat aan het einde van deze fase.

3.1. Fase 1: Literatuurstudie

Hier wordt er onderzoek gedaan naar reeds bestaande linters, al dan niet gerelateerd aan LaTeX. Er zal gekeken worden naar de programmeertalen waarin deze gemaakt zijn, hun werking, specificaties en functionaliteiten. Daarnaast

mag er ook niet vergeten worden om te kijken hoe een CI-pipeline effectief in zijn werk gaat, zodat het mogelijk is om de linter op een efficiënte manier te integreren in workflows. Indien er andere relevante zaken opduiken die belang kunnen hebben aan het onderzoek, zullen deze ook worden opgenomen om te onderzoeken. Op deze manier zal er geprobeerd worden een zo volledig mogelijk onderzoek te voeren. Een kijk naar hoe een linter *echt* werkt. Deze kennis zal gebruikt worden als inspiratiebron voor het uiteindelijke proof of concept dat opgesteld zal worden in een latere fase. Een persoonlijke deadline wordt op 1 mei 2024 geplaatst. Hoewel de focus op deze fase het grootst zal zijn gedurende de eerste twee weken, is het een fase die parallel blijft doorlopen gezien er altijd extra info nodig kan zijn doorheen de andere fasen. Het resultaat van deze fase is een (informeel) document waarin alle ondervindingen, onderzoeken en handige informatiebronnen bijgehouden zal worden over reeds bestaande linters en andere zaken die van pas komen tijdens zowel het onderzoeken als het ontwikkelen van deze proof of concept. Het document zal een bron van kennis zijn die gebruikt kan worden in verdere fasen.

3.2. Fase 2: Technische Analyse en Experimenten

In deze fase is het doel om de lijst van gewenste functionele en niet-functionele requirements aan te vullen en deze te structureren naar gelang hun prioriteit. Onder andere de keuze van de programmeertaal, eventuele libraries en andere softwaretools die gebruikt kunnen en zullen worden voor het maken en opzetten van de proof of concept zullen hier gebeuren. Een eerste poging rond een simpele CI-pipeline zal hier mogelijk ook plaatsvinden. Dit om de werking en configuratie ervan zo goed mogelijk te begrijpen. Fase 2 zal plaatsvinden gedurende de maand maart. Het resultaat is een overzicht van zowel de functionele als de niet-functionele requirements, op prioriteit gerankschikt. Alsook zal er een soort van sprint planning opgesteld worden om een overzicht te bewaren

3.3. Fase 3: Software Ontwikkeling (PoC - Proof of Concept)

In deze fase gebeurt het uitdagende werk. De proof of concept zal ontwikkeld worden. Een werkend prototype zal opgesteld worden, inclusief testen om de werking te garanderen en om het eenvoudig uitbreidbaar te maken. Niettemin zal er ook uitbundige documentatie geschreven worden, zodat alles wat er gebeurd duidelijk is voor elke vrijwilliger die een bijdrage wenst te leveren. Er wordt verwacht dat in deze fase enkel gefocust dient te worden op de effectieve creatie

van de linter met behulp van gevonden informatie en bronnen. Echt opzoekwerk zou minimaal moeten zijn in deze fase. Aan het einde van deze fase zal er dus een werkend prototype beschikbaar staan op een GitHub-repository van de auteur van dit onderzoek, Tristan Cuvelier. Er wordt verwacht dat deze fase wel enige tijd in beslag zal nemen gezien het verwachte resultaat en met een marge die ruim genoeg is voor problemen die opduiken tijdens het ontwikkelen van de software. Deze fase begint eind maart en duurt tot 17 mei.

3.4. Algemeen + Fase 4

Naast deze fasen is het ook nodig om de bachelorproef zelf uit te schrijven. Het is belangrijk dat dit doorheen de fasen heen ook al gebeurt. Zo kunnen de laatste paar weken gebruikt worden om alles op punt te zetten om de finale versie in te dienen. Deze fase start dus al gelijk met fase 1, op 19 februari waarbij er gelijdelijk aan doorheen de andere fasen aan het eind dossier gewerkt wordt. Deze fase heeft als deadline 21 mei gekregen, zo zijn er nog enkele reserve dagen voor de effectieve deadline van 24 mei om 12u 's middags. Het resultaat van deze fase is, zoals verwacht, een grondig uitgewerkte bachelorproef.

4. Verwacht resultaat, conclusie

Als resultaat voor deze bachelorproef wordt er verwacht dat er een open-source repository van een BibLaTeX-linter ter beschikking komt. Deze zal de studenten en lectoren van HoGent helpen om minder fouten te maken op vlak van bronvermeldingen binnen het schrijven van hun papers. Ook externe personen buiten HoGent zullen gebruik kunnen maken van deze linter. Hoewel de linter slechts een prototype zal zijn, is het de bedoeling dat deze als goede basis kan dienen om op relatief eenvoudige wijze aan verder te werken. Dit alles met als ultiem doel de ideale linter te worden voor BibLaTeX. Gezien dat er voorlopig slechts één BibLaTeX-linter online staat die niet snel werkend te krijgen is, zal dit ongetwijfeld een grote meerwaarde bieden aan iedereen die gebruik maakt van BibLaTeX.

Er wordt verwacht dat dit een uitdagende opdracht zal worden waaruit veel geleerd kan worden, 'trial and error' samen met grondig onderzoek zullen de basis zijn. Als softwaredeveloper-student is het een interessante uitdaging om zelf een linter te schrijven voor iets dat een student toegepaste informatica slechts enkele keren gebruikt doorheen de schoolcarrière. Pipelines zijn niet ongekend, maar blijven wel iets minder voorkomend in de tak van development in de richting.

Referenties

- Cassidy, J. (2013, juli). *Getting started with BibLaTeX*. Overleaf. Verkregen december 7, 2023, van https://www.overleaf.com/learn/latex/Articles/Getting_started_with_BibLaTeX
- Cuckow, P. (2022, december 17). *BibLaTeX Linter* (P. Cuckow, Red.). Verkregen december 7, 2023, van <https://github.com/Pezmc/BibLaTeX-Linter>
- Kamunya, T. (2023, januari 25). *Should You Start Linting Your Code? [+5 Linting Tools]*. Geekflare. Verkregen december 7, 2023, van <https://geekflare.com/linting-tools/#geekflare-toc-css-lint>
- Oetiker, T., Serwin, M., Partl, H., Hyna, I., & Schlegl, E. *The Not So Short Introduction to LaTeX: The Not So Short Introduction to LaTeX*. 2023, 19, 20. 298 p. Verkregen februari 13, 2024, van <https://tobi.oetiker.ch/lshort/lshort.pdf>