

nodigt U graag uit op de openbare verdediging van het proefschrift van

Maxime Jakubowski

ter behaling van de graad van Doctor in de wetenschappen

Gezamenlijk doctoraat met Universiteit Hasselt

Titel van het proefschrift:

**Shapes Constraint Language:
Formalization, Expressiveness, and Provenance**

Curriculum vitae

Promotors:

Prof. dr. Bart Bogaerts (VUB)

Prof. dr. Jan Van den Bussche (UHasselt)

De verdediging heeft plaats op

Vrijdag 31 mei 2024 om 16u

**UHasselt Campus Diepenbeek,
auditorium H1**

Samenstelling van de jury

Prof. dr. Marc Gyssens (UHasselt, voorzitter)

Prof. dr. Bas Ketsman (VUB, secretaris)

Prof. dr. Stijn Vansummeren (UHasselt)

Prof. dr. Anastasia Dimou (KULeuven)

Dr. Shqiponja Ahmetaj (TU Wien, Oostenrijk)

Dr. Slawek Staworko (RelationalAI, VS)

Maxime Jakubowski behaalde zijn Masterdiploma in Artificial Intelligence aan de Universiteit Maastricht in 2020. Na het afstuderen begon hij aan zijn doctoraat aan de UHasselt, binnen het Data Science Instituut onder begeleiding van Prof. Dr Jan Van den Bussche. Al snel werd het een gezamenlijk doctoraat met het AI Lab van de VUB onder begeleiding van Prof. Dr Bart Bogaerts. Deze samenwerking kwam tot stand door het gezamenlijk project van het Flanders AI Research Program waaronder dit doctoraat valt.

Het onderzoek valt op de intersectie van database theorie en artificieel intelligentie, waarbij formele logische technieken worden gebruikt om bestaande computertalen beter te begrijpen.

Abstract van het doctoraatsonderzoek

De Shapes Constraint Language (SHACL) is een schemataal voor het uitdrukken van structurele beperkingen op RDF-grafen, waarbij beperkingen op knopen “shapes” worden genoemd. We bestuderen verschillende aspecten van deze taal. Ten eerste tonen recente formalisaties een opvallende gelijkheid met description logics. We bouwen voort op deze formalisaties om SHACL als een logica te begrijpen. Bovendien, omdat de SHACL-specificatie alleen semantiek definieert voor niet-recursieve SHACL, zijn er enige inspanningen gedaan om recursieve SHACL-schema's mogelijk te maken. We betogen dat voor het definiëren en bestuderen van de semantiek van recursieve SHACL, inspiratie kan worden opgedaan uit onderzoek in non-monotonic reasoning. We bekijken de voorgestelde semantiek uit de literatuur en vergelijken deze met technieken uit goed gevestigd onderzoek naar non-monotonic reasoning.

Verder kunnen SHACL-expressies drie fundamentele constructies gebruiken die niet zo gangbaar zijn in vergelijkbare logica. Deze kenmerken zijn gelijkheidstests; disjunctheidstests; en closed. Het is niet duidelijk hoe de aanwezigheid van deze niet-standaard constructies de expressiviteit van SHACL beïnvloedt. We tonen aan dat elk van de drie constructies primitief is: met elke constructie kan men booleaanse queries uitdrukken die niet uitdrukbaar zijn zonder de constructie te gebruiken. We tonen ook aan dat de beperking die SHACL oplegt aan toegestane targets niet essentieel is, zolang closed niet gebruikt wordt. Daarnaast tonen we aan dat het verrijken van SHACL met “volledige” versies van gelijkheidstests of disjunctheidstests resulteert in een strikt krachtiger taal.

Tot slot stellen we een provenance semantiek voor voor SHACL. We stellen de notie voor van de “neighborhood” van een knoop dat een bepaalde shape in een graaf voldoet. Deze neighborhood is een subgraaf en biedt provenance voor de knoop en de gegeven shape. We stellen een correctheidseigenschap vast voor het verkregen provenance mechanisme, door te bewijzen dat neighborhoodsvoldoen aan de sufficiency eis die gearticuleerd is voor provenance semantiek voor databasequery's. We vergelijken neighborhoodsvoldoen met SPARQL-query's en bespreken implementatiestrategieën voor het berekenen van neighborhoodsvoldoen.