

De Online Geografie van Innovatieve Startups



GENEREREN VAN RUIMTELIJKE GEGEVENS OP BASIS VAN
WEBSITETEKSTEN OVER INNOVATIEVE BEDRIJFSACTIVITEITEN

Colofon

CONTACTGEGEVENS

Gegevens Student

Jaap Willems
Johannes de bekastraat 67
3514VL Utrecht
Studentnummer: 6003794
+316 11096362
willems_jaap@hotmail.com
jaap.willems3@students.uu.nl

Gegevens Stage



Innovatiespotter
Protonstraat 3A
9743 AL Groningen
050 3134898
06 36511009
info@innovatiespotter.nl

Gegevens Studie



Universiteit Utrecht
Human Geography
Faculteit Geowetenschappen
Heidelberglaan 6, 3584 CS
Postbus 80125, 3508 TC Utrecht

Gegevens begeleider

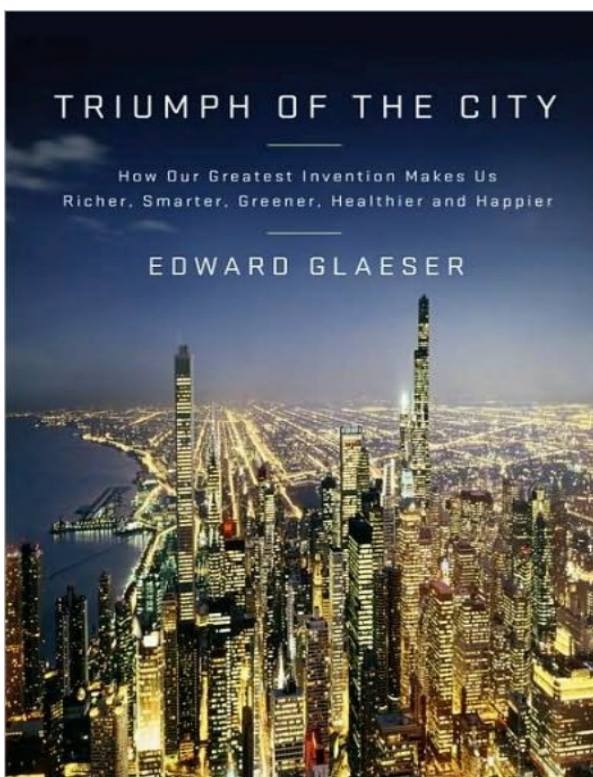
M.J. Smit
M.J.Smit@uu.nl

Inleverdatum:

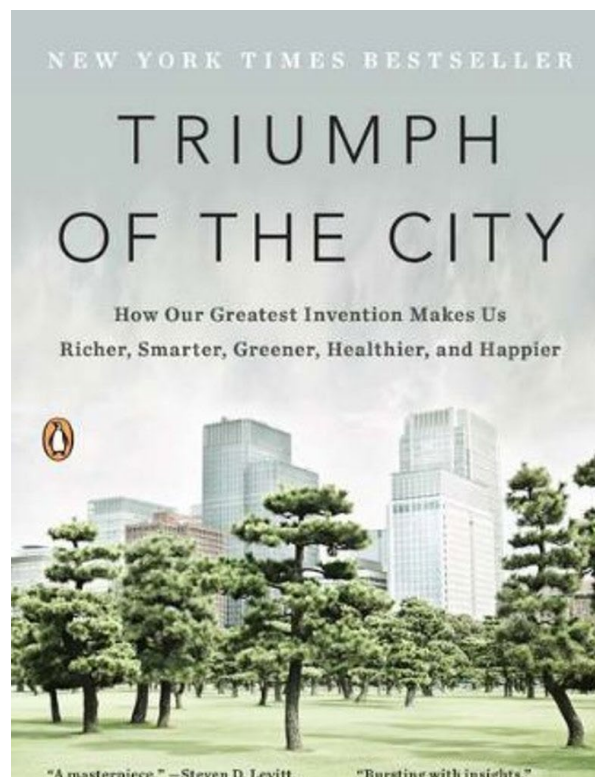
8-3-2019

Formaat:

A4 PDF



Figuur 0.1 Boekomslag van Edward Glaeser's boek: *Triumph of the City*. Bron: Glaeser (2011)



Figuur 0.2 Boekomslag van Edward Glaeser's boek: *Triumph of the City*. Bron: Glaeser (2011)

Voorwoord

In 2017 begon mijn master Human Geography met de introductie cursus Urban Futures over de sociaal-ruimtelijke en economische ontwikkelingen in stedelijke gebieden. Centraal in deze cursus stond het boek Triumph of the City van Harvard professor Edward Glaeser (2011). Als student met een achtergrond in stedelijk ontwerp en sociale thema's had ik nog nooit van Glaeser gehoord, en ook niet veel over economische theorieën laat staan van ruimtelijke econometrie.

Hoe anders is dit nu.

Voor mijn gevoel moest ik tijdens mijn scriptie iedere keer op nul beginnen. Ik weet niet wie of wat mij er doorheen heeft gesleept, maar toevallig duikt Glaeser gedurende mijn master overal op. Tijdens mijn fieldtrip naar Chicago heb ik hem overigens niet gezien.

Zijn ideeën over innovatiekracht en verticale verstedelijking vond ik eigenlijk best interessant. Volgens Glaeser is de beste manier om te floreren om een grotere diversiteit aan talent en economische activiteiten in dichtbevolkte gebieden te faciliteren. Zoals op de linker boekomslog te zien, vrij radicaal. Die rechter valt wel mee.

Glaeser krijgt een plekje in mijn voorwoord.

Met deze scriptie komt een eind aan mijn master stadsgeografie. Veel lees en kijk-plezier,

Jaap Willems

Samenvatting

Ruimtelijke processen van kennisuitwisseling en startup-formatie spelen een cruciale rol bij innovatie als motor van economische groei. De huidige praktijk waarin innovatieve bedrijfsactiviteiten worden vastgesteld, zijn vaak gebaseerd op patentaanvragen, verkregen durfkapitaal, R&D-indicatoren, vragenlijsten of gereaaliseerde groei. Zelfs wanneer het voorgaande samen wordt genomen, blijven er aanzienlijke beperkingen bestaan. Deze beperkingen betreffen vooral de kleinste bedrijven zoals startups. Dit onderzoek maakt gebruik van een relatief nieuwe en nog weinig toegepaste benadering in de economische geografie om innovatieve activiteiten van startups te ramen op basis van de teksten op hun websites.

Dit werkt als volgt: de tekst van geïndexeerde bedrijfswebsites wordt onttrokken door een geautomatiseerde *webscraper*. Deze tekst is vervolgens aan algoritmes van de 'Innovatiespotter' onderworpen en op basis van een classificatiemodel gerangschikt op relevantie uitgesplitst naar tien innovatieve thema's gebaseerd op de Nederlandse topsectoren. De ruwe resultaten zijn op basis van gekoppelde inschrijfgegevens gefilterd met als doel dat alleen individuele kleine en jonge bedrijven die classificeren als startup over zouden blijven.

Deze procedure heeft geresulteerd in een dataset die is gebruikt voor verdere ruimtelijke economische analyse in Noord-Nederland. Doormiddel van een gerichte steekproef is de samenstelling en kwaliteit van de resultaten gecontroleerd. Met zowel het valideren van de methode als het analyseren van de resultaten, komt deze studie tegemoet aan de uitdaging zoals geagendeerd door bibliografisch onderzoek over het ondernemen ecosysteem om potenties van verschillende innovatieve omgevingen beter aan het licht te brengen.

Trefwoorden: webscraping, automatische classificatie, topsectoren, innovatie, ondernemerschap, ondernemend ecosysteem, micro-geografie, startups, bedrijfslocatie

Abstract

Spatial processes of knowledge exchange and start-up formation play a crucial role in innovation as a driver of economic growth. Current practice in which innovative business activities are identified are often based on patenting, obtained venture capital, R&D indicators, questionnaires about innovative activities or realized growth. Even when the foregoing is taken together, considerable limitations remain. These limitations mainly concern the smaller companies such as startups. This research uses a relatively new and rarely applied approach in economic geography to estimate innovative activities of startups based on the texts on their websites.

The process can be described as follows: the text of indexed company websites is extracted by an automated web scraper. This text was then subjected to the algorithms of the 'Innovatiespotter' based on a classification model for ten innovative themes based on the Dutch topsectors. The scores were then ranked by relevance. Subsequently, the raw results including linked registration data were filtered so that only individual small and young companies that would classify as a startup would remain.

This procedure has resulted in data that has been used for further spatial economic analysis in the Northern Netherlands. The composition and quality of the results was checked by means of a targeted sample. With the validation of the method and analysis of the results, this study meets the challenge as placed on the agenda by bibliographical research on the entrepreneurial ecosystem to further expose the potential of different innovative environments.

Keywords: webscraping, automatic classification, Dutch topsectors, innovation, entrepreneurship, entrepreneurial ecosystem, micro-geography, startups, business location

Inhoud

	Colofon	3
	Voorwoord	5
	Samenvatting	6
	Abstract	7
H 1	Startups in het Innovatielandschap	10
	1.1 Aanleiding	10
	1.2 Probleemstelling	10
	1.3 Doelstelling	11
	1.4 Vraagstelling	11
	1.5 Geografische Afbakening	12
H 2	Theoretisch Kader	13
	2.1 Meten van Innovatie	13
	2.1.1 Traditionele indicatoren	13
	2.1.2 Vragenlijsten	14
	2.1.3 Media- en literatuuranalyse	15
	2.1.4 Gecombineerde databronnen	16
	2.1.5 Big data	17
	2.1.6 Webscraping: pioniers in de geografie	18
	2.1.7 Webscraping: positieve eerste resultaten	20
	2.1.8 Webscraping: toepassing per thema	21
	2.2 Meten van Startups	23
	2.2.1 Startup definities	23
	2.2.2 Opstartdatum	23
	2.2.3 Nieuwheid	24
	2.2.4 Dynamische activiteiten	25
	2.2.5 Het belang van onafhankelijkheid	25
	2.3 Het Ondernemende Ecosysteem	29
	2.3.1 Nationale identificatie, regionaal speelveld, wereldwijde banden	29
	2.3.2 Kennisontwikkeling en het sub-urbane landschap	33
	2.3.3 Verklaringen voor startup-formatie	37
H 3	Methode	42
	3.1 Onderzoekspopulatie en Steekproef	42
	3.1.1 Topsector-expertise en innovativiteit	42
	3.1.2 Startup-kenmerken	44
	3.1.3 Onafhankelijkheid en herkomst	45
	3.2 Webscraping	46

3.2.1	Procesbeschrijving	47
3.2.2	Filteren van de inschrijfgegevens	48
3.2.3	Selectie op relevantiescore	51
3.3	Ruimtelijke analyse	53
3.3.1	Samenvoegen, geocoderen en koppelen van externe data	53
3.3.2	(On)afhankelijke variabelen	54
3.3.3	Ruimtelijke verklarende statistiek	56
H 4	Resultaten	58
4.1	Steekproef-controle	58
4.2	Geografische Analyse	63
4.2.1	Beschrijvende statistiek	63
4.2.1	Verklarende statistiek	64
4.2.1a	Vergelijking op innovativiteit	65
4.2.1b	Vergelijking op thema's	67
4.2.1c	Overige afwijkingen	70
H 5	Conclusie & Discussie	72
5.1	Conclusie	72
5.1.1	Identificatie innovatieve startups	72
5.1.2	Ruimtelijke en thematische inzichten	74
5.2	Discussie	76
5.2.1	Sterke en zwakke punten	76
5.2.1a	Webscraping en automatische classificatie	76
5.2.1b	Ruimtelijke technieken en resultaten	78
	Referenties	80
H B1	Steekproef-Controle	90
H B2	Opschonen Externe Data	122
H B3	Overige Figuren en Tabellen	124
B3.1	Variabele Kennisinstelling	124
B3.2	Figuren t.b.v. de poissonverdeling	125
B3.3	Thematische kaarten	126
H B4	Ruimtelijke Statistiek	128
B4.1	Ruimtelijke Autocorrelatie	128
B.4.4.1	Conceptualisatie ruimtelijke relaties	129
B.4.4.2	Global Moran's Toets	131
B.4.4.3	Modelleren van de verstoringen	132

1 Startups in het Innovatielandschap

In Edward Glaeser's (2011) boek 'Triumph of the City' wordt het leven in hogere dichtheden bestempeld als enige juiste keuze voor een welvarende en milieuvriendelijke toekomst. Steden zijn hierin ontwikkelingscentra: combinaties van dichtheid en menselijke creativiteit die nieuwe ideeën en producten vormen. Dit centrale idee wordt ondersteund door statistieken, waarmee wordt aangetoond dat clusters met hoger opgeleide mensen zich steeds opnieuw uitvinden en op die manier sneller aanpassen aan een veranderende omgeving. Met als onderliggende idee dat een moderne economie niet zonder goede ideeën kan, wordt een sleutelrol toegewezen aan startups als motor van de innovatieprocessen (Colombo & Piva., 2008; Davila et al., 2003; Mustar et al., 2008; Europese Commissie, 2013; Pan & Yang, 2018). Mede daarom is er veel aandacht voor de dynamiek in het innovatielandschap, en in het bijzonder voor innovatieve startups. Dit geeft aanleiding om vanuit een geografische invalshoek nader te kijken naar innovatieve startup-activiteiten.

1.1 Aanleiding

Uit een recent overzicht van de literatuur blijkt dat de wetenschappelijke aandacht voor innovatieve omgevingen alleen maar is toegenomen (Malecki, 2018). Beleidsmakers en economische theorieën ondersteunen clustering van kennis ten behoeve van innovatie en productiviteit. Toch is er kritiek en bestaat er onduidelijkheid in het publieke debat over wanneer, hoe en of clusters wel actief moeten worden gestimuleerd (zie kader 1.1). Ondanks aansprekende en succesvolle voorbeelden van innovatieve omgevingen, blijkt ook in het wetenschappelijke debat nauwelijks consensus te bestaan over de rol die de omgeving precies speelt of zou kunnen spelen (Malecki, 2018). Het is bovendien onduidelijk wanneer welk concept bruikbaar is voor welke regio. Wel wordt door zowel de wetenschap als het bedrijfsleven bevestigd dat er economische voordelen verbonden zijn aan geografische concentraties van onderling verbonden bedrijven en instellingen.

Volgens het Planbureau voor de Leefomgeving (2017) is in succesvolle steden en regio's het beleid de laatste decennia opgeschoven van fysiek beleid (gebouwde omgeving) naar beleid gericht op een totaal (innovatie)systeem. Zowel voor het actief stimuleren van fysieke clusters, als voor het sturing geven aan ecosystemen, blijft echter de

behoefte bestaan om ingrepen te baseren op de juiste informatie. Meer kennis over startups in het innovatielandschap zou daarom een betere basis vormen voor verdere ontwikkelingen en het daarbij behorende economische beleid. Innovatieve startups spelen een cruciale rol binnen de dynamiek van het innovatielandschap, maar zijn soms lastig op te sporen. Om succesvolle innovatiemilieus te stimuleren is het echter wel van belang om juist deze groep beter in beeld te kunnen krijgen.

1.2 Probleemstelling

Een van de uitdagingen die nog op de onderzoeksagenda staat, is het beter in beeld brengen van innovatieve startups in verschillende regio's (Malecki, 2018). Onderzoek door economische geografen zou vanuit een geografische context de potentie van innovatieve omgevingen nog beter aan het licht moeten brengen. Dergelijke studies zouden ondernemers en beleidsmakers helpen met het integreren van een ondernemend ecosysteem concept in regionale ontwikkeling. Een belangrijk onderdeel hiervan is een goed beeld van de startups verkrijgen om deze vervolgens in hun behoefte te kunnen voorzien. Zo blijkt bijvoorbeeld uit onderzoek dat er vanuit startups in de regio Utrecht behoefte is aan meer persoonlijke betrokkenheid vanuit de publieke leiders in de re-

Kader 1.1 Debat over het ideale innovatielandschap

Het meeste bekende archetype van een startup- en innovatielandschap is een innovatiecluster. Volgens de opvattingen van Glaeser (2003) zou een diversiteit aan kennis in ‘hoge’ – letterlijk en figuurlijk – dichtheden centraal moeten staan. Innovatieclusters gaan over specialistische kennis binnen een toereikende afstand om ontmoeting mogelijk te maken. Silicon Valley is het meest besproken voorbeeld van een succesvol innovatiecluster, maar van diversiteit en hoge dichtheid is weinig sprake. Glaeser speculeert daarom dat Silicon Valley het op termijn wel eens moeilijk zou kunnen krijgen. Zo heeft de geschiedenis bewezen dat te veel specialisatie in lage dichtheden op den duur fataal kan worden. Moeilijkheden ontstaan zodra gevraagd wordt mee te bewegen met de tijd. Nieuwe ontwikkelingen vereisen aanpassing en flexibiliteit. Volgens Glaeser zijn diversere omgevingen dynamischer, en daarom beter in staat zich aan te passen aan veranderingen. Glaeser nuanceert deze visie door te wijzen op gunstige intrinsieke kenmerken (*‘entrepreneurial mindset’*) en het grote aantal bedrijven die aanwezig zijn in Silicon Valley. Laatstgenoemde voorwaarden verhogen de dynamiek en bevorderen daarmee ook de toekomstbestendigheid van de regio. Vanuit deze optiek is enige variatie mogelijk, maar blijft het vooralsnog onduidelijk – zelfs vanuit voorstanders van clustering – op welke manieren en onder welke voorwaarden kennis het beste geografisch verspreid en verbonden zou moeten worden om een optimaal innovatiemilieu te laten ontstaan.

In navolging van het voorgaande brengt het debat ook minder optimistische geluiden voort betreft de inzet op innovatieclusters. Zo neemt sociaal geograaf Smit (2010) duidelijk stelling over beleid dat wetenschapsparken stimuleert: ‘Men stapelt science park op science park zonder dat de toegevoegde waarde van deze parken is bewezen. Die is er naar mijn mening niet, en de agglomeratie-effecten worden dan ook schromelijk overschat’ (Smit, 2010). Een recenter voorbeeld van kritiek komt uit de rubriek ‘de ondernemer’ van het Algemeen Dagblad. Hierin schrijft Dries Faems, hoogleraar innovatie & organisatie, dat samenwerking met externe partners om sneller en effectiever nieuwe producten op de markt te brengen te vaak en te makkelijk wordt gezien als wondermiddel (Faems, 2017): ‘Wat ondernemers vaak uit het oog verliezen is dat open innovatie ook risico’s en kosten met zich meebrengt’. Ron Boschma, hoogleraar regionale economie en innovatiestudies, gaat nog een stap verder door te stellen dat er te weinig sprake is van een kritisch debat over innovatieclusters. Volgens hem komt dit doordat bijna iedereen wel bij een innovatieproject betrokken is en door een geloof in clusters die vooral gebaseerd is op anekdotes, bijvoorbeeld over Silicon Valley. Volgens hem ontbreekt het aan empirisch onderzoek naar resultaten van clusters omdat evaluatie niet hoog op de agenda staat (Zeemeijer, 2016).

gio (EBU Research, 2016) en uit onderzoek in de regio Noord-Nederland blijkt dat ondernemers een gebrek aan kennis zien als obstakel om subsidie aan te vragen ten behoeve van innovatie (Faems & Wennink, 2018). Ook de internationale literatuur onderkent dat startups vanwege hun geringe omvang lijden aan een structureel gebrek aan zowel tastbare als ontastbare hulpbronnen (Wymer & Regan, 2005).

1.3 Doelstelling

Het doel van dit onderzoek is om ruimtelijke gegevens over innovatieve startup-activiteiten te genereren door middel van een efficiënte *web-based* methode, en daarmee nieuwe inzichten te bieden in het innovatielandschap. Specifiek gaat het om de innovatieve startups in Noord-Nederland. Er zijn daarmee twee onderzoeksdoelen: de *web-based* methode valideren en het resultaat van die methode analyseren. Er kan worden gesproken van een

exploratief en hoofdzakelijk kwantitatief onderzoek.

1.4 Vraagstelling

De hoofdvraag die centraal staat in deze scriptie is:

In hoeverre kan het aantal innovatieve startups in Noord-Nederland in beeld worden gebracht met behulp van een web-based zoekmethode en welke geografische en thematische inzichten levert dat op?

Deelvragen:

1. *Hoe kunnen innovatieve startups worden geïdentificeerd en welke classificaties zijn geschikt ten behoeve van een kritische evaluatie in deze studie?*
2. *Vanuit welke geografische theorieën kan de opkomst en ontwikkeling van innovatieve startups worden verklaard?*
3. *Waar zijn de innovatieve startups in het onder-*

zoeksgebied gevestigd?

4. In hoeverre relateren de gevonden startups aan de verwachtingen op basis van het theoretisch kader en welke aanvullende empirische en methodologische inzichten kunnen worden verkregen?

1.5 Geografische Afbakening

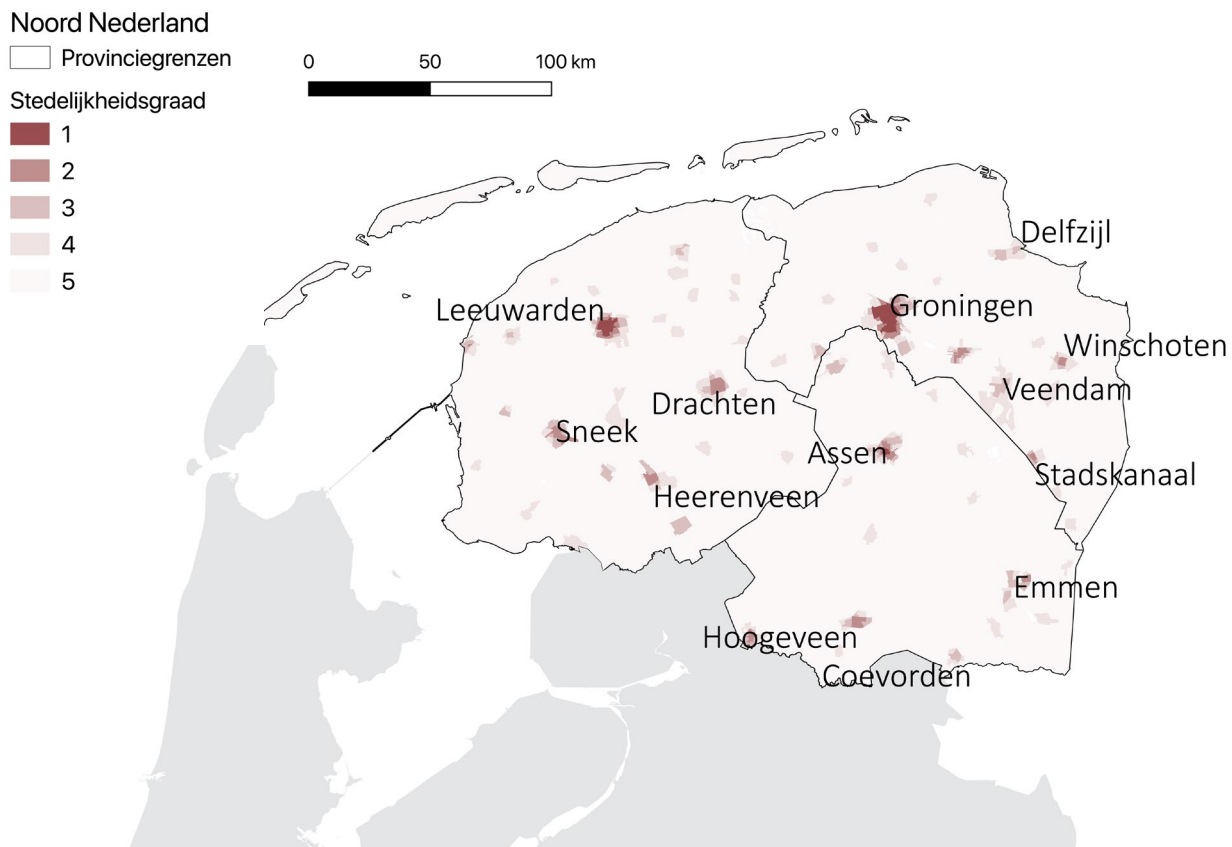
Als onderzoeksgebied (Figuur 1.1) is gekozen voor de regio Noord-Nederland bestaande uit de provincies Groningen, Friesland en Drenthe. Het gekozen schaalniveau leidt tot een werkbaar hoeveelheid gegevens in overeenstemming met de omvang van deze studie. Daarnaast is er relatief weinig onderzoek op bedrijfsniveau gedaan naar dit schaalniveau en type ecosysteem (periferie). Er wordt bijvoorbeeld meer onderzoek gedaan in en rondom de Randstad. Ten slotte zijn de beschikbare gegevens ook beter geschikt voor de analyse van Noord-Nederland, aangezien relatief veel websites voor deze regio zijn geïndexeerd (zie methodeparaagraaf 3.2 met tabel 3.2).

1.6 Leeswijzer

Dit onderzoek is als volgt gestructureerd. Na dit inductie hoofdstuk wordt in het theoretisch kader (H2) besproken wat de voorwaarden moeten zijn die bijdragen aan een succesvolle meting van inno-

vatie en startups. Ter afsluiting volgt een overzicht van geografisch relevante theorieën en empirische resultaten van studies die kunnen helpen bij het verklaren van de ruimtelijke startup-formatie in Noord-Nederland. Antwoord op de eerste twee deelvragen zullen voornamelijk gebaseerd zijn op de bevindingen uit het theoretisch kader. In het daaropvolgende hoofdstuk wordt de methode (H3) die ten grondslag ligt aan de ruimtelijke analyse behandeld. Dit is de *web-based* methode *webscraping*. Bij het bespreken van de verzameling en verwerking van de data worden ook de analysetechnieken verantwoord. Vervolgens worden de resultaten (H4) in tekst, tabel en beeld gepresenteerd. Het resultaten hoofdstuk heeft vooral betrekking op deelvraag drie, maar ook op de vierde deelvraag. De conclusie (H5) sluit af met antwoord op de hoofd- en deelvragen, en eindigt met een discussie en aanbevelingen.

De hoofdstukken zijn onderverdeeld in deelhoofdstukken en paragrafen. Naast genummerde tabellen en figuren is in sommige paragrafen in hoofdstuk twee gebruik gemaakt van aanvullende tekstkaders. Hier wordt in de tekst naar verwezen voor relevante achtergrondinformatie die mogelijk interessant is voor de lezer, maar niet noodzakelijk is om de hoofdtekst te begrijpen.



Figuur 1.1 Onderzoeksgebied Noord-Nederland naar stedelijkheidsgraad. Bronnen: CBS (2018) en het Kadaster (2018).

2 Theoretisch Kader

Dit hoofdstuk bespreekt de theoretische basis voor het meten van innovatie, en in het bijzonder het potentieel aan innovatie vanuit de startups in een regio. Eerst wordt in deelhoofdstuk 2.1 uitgebreid ingegaan op het begrip innovatie en hoe dit kan worden geoperationaliseerd ten behoeve van regionaal onderzoek. Hierbij is een belangrijk punt van aandacht dat een startup als relatief klein en nieuw bedrijf gevonden moet kunnen worden. Deelhoofdstuk 2.2 gaat dieper in op het begrip startup en de operationalisering hiervan. Deelhoofdstuk 2.3 sluit af met de geografische context waarbinnen de opkomst van innovatieve startups geanalyseerd en verklaard kan worden.

2.1 Meten van Innovatie

Ondanks het feit dat innovatie al in 1934 door Schumpeter breed werd gedefinieerd als introductie van een nieuw product, dienst, productiemethode, organisatievorm, opening van een nieuwe markt of het aanboren van nieuwe bronnen (Schumpeter & Backhaus, 2003), zijn statistieken die vóór de jaren zeventig innovatie beoogde te meten bijna uitsluitend gericht op technologische innovatie (Godin, 2002). Met de term innovatie werd en wordt vaak technologische innovatie bedoeld zonder dit expliciet als zodanig te benoemen (Godin, 2008).

De tweedeling die tegenwoordig gemaakt wordt op bedrijfsniveau omvat naast technische innovaties bestaande uit nieuwe technologieën, producten en diensten, ook de sociale of administratieve innovaties bestaande uit nieuwe procedures, beleid en organisatievormen (Van de Ven & Poole, 1990; Garcia & Calantone, 2002). Dit neemt niet weg dat de technologische nadruk nog steeds de overhand heeft, dit blijkt bijvoorbeeld uit het Nederlandse innovatiebeleid gericht op topsectoren (tabel 2.1) die overwegend technologisch van aard zijn. En relatief weinig onderzoek naar innovaties in de creatieve industrie (Camelo-Ordaz, Fernández-Alles, Ruiz-Navarro, & Sousa-Ginel, 2012).

2.1.1 Traditionele indicatoren

Afgezien van de technologische nadruk, zorgen aanvullende problemen van traditionele methoden zoals het gebruik van R&D indicatoren voor

een beperkt beeld van innovatie. Dergelijke traditionele methoden worden nog steeds gebruikt, met als gevolg dat ook de meetproblemen nog steeds bestaan. Wel constateert Godin (2002) dat er sinds eind vorige eeuw progressie is geboekt gezien de conceptuele verschuiving van overwegend resultaat meten (output) naar investeringen en activiteiten meten (input). Innovatie wordt echter gezien als een multidimensionaal fenomeen, en daarom zijn zowel in- als output belangrijk voor het meten van innovatie. Dit onderzoek is gericht op het innovatie- en groeipotentieel van startups en zal daarom de nadruk leggen op de theorie betreffende het meten van de input, maar wel rekening houdend met het feit dat dit niet los van de output kan worden gezien.¹

Een complete meting van zowel in- als output is juist bij een veelzijdig concept als innovatie niet gemakkelijk. Wat betreft de input wordt doorgaans gekeken naar de innovatieve activiteiten van de ondernemingen. Echter kunnen onder inputmetingen ook de investeringskosten vallen (Godin, 2002) terwijl andere onderzoeken investeringen juist duiden als een eerste resultaat en dus als een meting van de output (Fazio, Guzman, Murray, & Stern, 2016). Input van innovatie meten op basis van de activiteiten werd van begin af aan gezien als

¹ Ter verduidelijking; minder relevante resultaatmetingen zoals specifiek patenten, risicokapitaal en omzet zullen niet uitgebreid worden besproken.

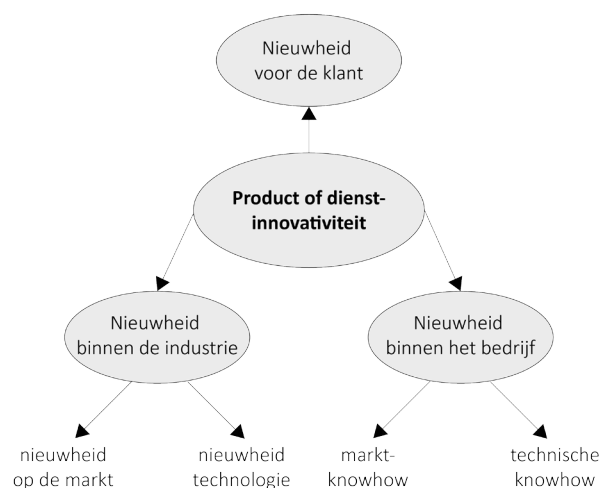
Tabel 2.1 Topsectoren en kenmerkende activiteiten.

Topsector	Kenmerkende activiteiten en of -toepassingsgebieden
Agri & Food	Diverse (plantaardige en dierlijke) voedselketens: consument en maatschappij, klimaatneutraal en circulair voedsel, gezond en veilig voedsel, slimme technologie
Chemie	Petrochemie, basischemie, elektrochemie en fijnchemie.
Creatieve Industrie	Architectuur, gaming, human empowerment, industrieel ontwerp, media, kunst en innovatie, persoonlijke ervaring, strategie, waardecreatie
Energie	Verduurzaming van de energiehuishouding: wind op zee, energie en industrie, maatschappelijk verantwoorde energietransitie, digitalisering, nieuw gas, biobased economy, internationale energie export, urban energy, systeemintegratie, human capital agenda
High Tech Systemen en Materialen	Advanced instrumentation, aeronautics, automotive, electronics, embedded systems, healthcare, hightech materials, lighting, nanotechnologie, photonics, cyber security, advanced printing, semiconductor equipment, smart industry, space, staal
Life Science & Health	Volks- en diergezondheid: Vaccins, diagnostica, farmaca, biomedische materialen, preventieve technieken en middelen.
Logistiek	Internationale supply chains, regierol op knooppunten, servicelogistiek, innovatie luchtvaart, goederenvervoer water
Tuinbouw en uitgangsmaterialen	Veredeling van groenten, fruit, bomen, bloemen & bollen, groene leefomgeving, plantgezondheid, fytosanitairste ketens, greenhouse technology, precisietuinbouw, digitale transformatie, energie en CO2
Water	Water- en delta technologie, maritieme bouw, water als resource, waterzuivering

Bronnen: Verhagen, 2011 en Rijksoverheid, 2019

lastig (Godin, 2002). Bij dergelijke pogingen wordt aan bedrijven gevraagd hoeveel ze uitgeven aan innovatie (vaak R&D) of worden projecten geanalyseerd op basis van innovatieve activiteiten. Beide manieren zijn in de eerste plaats ontoereikend omdat veel bedrijven geen gegevens bijhouden over hun eigen innovatie of deze niet willen delen. Daarnaast ontstaan aanvullende problemen afhankelijk van de gebruikte methode om de informatie te verkrijgen of om de innovatieve bedrijven op te sporen. Problemen die optreden hebben te maken met de onderzoekspopulatie, selectiecriteria en de respons.²

² Zie (Godin, 2002, p. 22; van der Panne, 2007, p. 496-498)



Figuur 2.1 Operationalisering van innovativiteit. Bron: Garcia & Calantone (2002)

Het grootste meetprobleem heeft volgens Godin (2002) echter te maken met het concept van innovatie zelf. Innovatie is een vaag concept, waarbij het lang niet altijd duidelijk is of iets nieuw is op wereldniveau, in eigen land of vooral vanuit de perceptie van een onderneming en in een bepaalde industrie. In studies van Garcia & Calantone (2002) en Smith (2005) wordt kritisch gekeken naar alle verschillende operationalisering van innovatie. Ook hier wordt geconcludeerd dat voor wie of wat iets nieuw is, cruciaal is voor de bepaling of iets innovatief is (figuur 2.1). Er wordt benadrukt dat innovatie per definitie iets kwalitatief nieuws is waardoor het lastig te kwantificeren is. Toch zijn er verschillende methoden die het concept van innovatie op zijn minst dicht benaderen, hiervan moet bepaald worden welke het beste in aanmerking komen als basis voor dit innovatie onderzoek.

2.1.2 Vragenlijsten

De *Community Innovation Surveys* (CIS) zijn een reeks onderzoeken die worden uitgevoerd door nationale bureaus voor de statistiek in heel Europa (Centraal Bureau voor de Statistiek [CBS], 2016; Eurostat, 2017). Het gaat hierbij om een wereldwijd veel gebruikte methode: steekproefsgewijs worden bedrijven benaderd om een vragenlijst in te vullen. Meestal is het hoofddoel om innovatiestatistieken te verzamelen voor de gehele onderzoekspopulatie. In het geval van CIS gaat het dan om innovatiestatistieken van bedrijven in een

regio. De CIS statistieken geven inzicht in de product-, proces-, organisatie- en marketinginnovaties bij bedrijven (CBS, zd) en kunnen een aanvulling zijn op traditionele R&D gegevens.

Aanvulling op traditionele gegevens blijkt nodig, want ondanks de zorgvuldige samenstelling van R&D-cijfers geven deze geen volledig beeld van alle innovatieve activiteiten (van der Panne, 2007). Bovendien zijn ook de aanvullingen van simpelere R&D data niet toereikend. Zo zouden vooral kleine bedrijven de neiging hebben om hun kleinschalige en informele R&D-activiteiten niet volledig te rapporteren (Kleinknecht, Van Montfort, & Brouwer, 2002). Sterker nog, CIS-gegevens uit Nederland geven uitsluitend een beeld van bedrijven met minimaal 10 werkzame personen, en brengen dus niet de innovatieve activiteiten van de kleinste bedrijven (micro-schaalniveau) in beeld terwijl deze groep wel potentieel interessante startups kan bevatten³ (CBS, 2016). Ook bestaan er vertekeningen op basis van de sector waarbinnen men innovatie onderzoekt. Kleinknecht et al. (2002) constateren een systematische voorkeur voor de verwerkende industrie, terwijl binnen de dienstsector de neiging bestaat innovatie te onderschatten. Overige tekortkomingen van de R&D indicator hebben te maken met eerdergenoemde problemen zoals selectiecriteria, respons, en het feit dat input los van output wordt gemeten, terwijl beide indicatoren nodig zijn.

2.1.3 Media- en literatuuranalyse

Van der Panne (2007) stelt dat veel van de bovengenoemde nadelen niet van toepassing zijn bij de 'op literatuur gebaseerde innovatie-output' (LBIO)-methode. Deze methode screent gespecialiseerde vakbladen op nieuwe productaankondigingen en wordt door van der Panne afgezet tegenover het gebruik van CIS-gegevens. Het gaat bij LBIO om eerste resultaatmetingen; alleen bedrijven die met succes hun innovatieve activiteiten (input) omzetten in nieuwe producten (output) worden geïdentificeerd. De LBIO methode is efficiënt aangezien bedrijven niet individueel hoeven te worden ondervraagd. Dit heeft als voordeel dat innovaties van kleine bedrijven wel kunnen worden meegenomen, terwijl CIS- en R&D enquêtes om redenen van efficiëntie meestal een minimale bedrijfsomvang hanteren. Zonder individuele bedrijven te belas-

³ (Hollanders, Es-Sadki, & Kaner, 2016) Van de deelnemende EU landen blijkt Nederland het minst vaak CIS-enquêtes af te nemen, Bovendien wordt er niet 'bijgeschat' voor bedrijven op het microschaalniveau.

ten, vermijdt de LBIO methode problemen van geheimhouding en non-respons. Volgens Flor & Oltra (2004) kunnen LBIO-gegevens mede daarom worden beschouwd als meest omvattende en uitgebreide vorm van secundair data gebruik voor innovatieonderzoek.

Ondanks het feit dat met LBIO-gegevens kleine bedrijven geïdentificeerd worden, hebben juist kleine bedrijven niet altijd de middelen om productaankondigingen onder te brengen in een geschikt tijdschrift. Kleine bedrijven blijven daardoor ondervertegenwoordigd. Van der Panne (2007, p. 498) verwijst verder naar verschillende onderzoeken die aanwijzingen vinden dat factoren zoals de sector en bedrijfsgrootte in verschillende landen verschillend tot uiting komen, en daarmee de representativiteit van de LBIO-gegevens beïnvloeden. Dit betekent dat er kritisch gekeken moet worden naar de beweegredenen die schuilen achter de aankondigingen van producten in specifieke tijdschriften. Selectie of valse tellingen kunnen al snel leiden tot een vertekend beeld.

Uit de analyse van Van der Panne (2007) zelf, blijkt dat 60 procent van de nieuwe productaankondigingen slechts dient om distributie van een product verder op gang te helpen in plaats van het verder ontwikkelen of aankondigen van een nieuwe innovatie. Dit heeft tot gevolg dat de innovatie in het binnenland ernstig wordt overschat omdat een buitenlandse innovatie in het binnenland als nieuw wordt gepresenteerd. De algehele conclusie die in het vergelijkend onderzoek door van der Panne wordt getrokken, is dat het gebruik van LBIO-gegevens wel degelijk geschikt is om innovatie te meten. Tegelijkertijd schiet deze methode soms tekort wat betreft het scheppen van een representatief regionaal beeld van de daadwerkelijke innovatie. Voor een nauwkeurigere samenstelling van LBIO-gegevens wordt daarom gesuggereerd om de origine van elke aankondiging van een innovatie te verifiëren, en dat gaat weer ten kosten van de efficiëntie.

Toch blijkt er voor identificatie van innovatieve startups – op basis van het vergelijkende onderzoek door van der Panne – nog een aanzienlijk voordeel te bestaan. Want ondanks het feit dat bij LBIO-gegevens kleinere bedrijven nog steeds ondervertegenwoordigd zijn, worden ze wel gevonden; maar liefst 62 procent van de gevonden innovaties zijn afkomstig van servicebedrijven en bedrijven met minder dan 10 werknemers. LBIO-

gegevens lijken daarmee effectief in het meten van innovatie in servicebedrijven en kleine bedrijven, en daarmee dus ook relatief effectief voor het in beeld brengen van innovatieve startups.

2.1.4 Gecombineerde databronnen

Uit voorgaande paragrafen blijkt dat R&D- en CIS-gegevens minder geschikt zijn voor identificatie van de kleinste bedrijven. Maar in tegenstelling tot LBIO-gegevens zijn R&D- en CIS-gegevens wel meer gericht op de innovatie input, terwijl LBIO-gegevens slechts de resultaten van innovatie identificeren. In de wetenschappelijk literatuur is relatief weinig kwantitatief onderzoek te vinden gericht op het volledig in beeld brengen van innovatieve bedrijven op het micro schaalniveau. Daarentegen lijken overheden en particulieren bedrijven zich hier meer mee bezig te houden.⁴ De benodigde microdata is meestal afkomstig van de nationale overheidsinstelling voor de statistiek in combinatie met een bedrijvenregister⁵ en eventueel aanvullende datasets die door middel van vragenlijsten tot stand komen (Jarmin & Miranda, 2002; Kloek, 2002). Er circuleren meerdere (internationale) rapporten op het internet die innovatie of het innovatiepotentieel indexeren gebruikmakend van verschillende datasets die door nationale statistische bureaus of private onderzoeksbureaus zijn samengesteld. In tabel 2.2 zijn enkele voorbeelden opgenomen van recente publicaties die op grotere schaal een vorm

van innovatie potentieel indexeren.

De rapporten in tabel 2.2 gebruiken verschillende technieken om het innovatiepotentieel – inclusief dat van de kleinere bedrijven – vast te stellen. In sommige gevallen wordt er enkel gekeken naar contextuele factoren die van invloed zijn op de manier waarop bedrijven functioneren. Indicatoren hiervan zijn bijvoorbeeld de aanwezigheid van kennis of gunstige bedrijfsregels in de regio waardoor het makkelijk is om een bedrijf te starten. Deze indicatoren helpen niet bij het opsporen van kleine innovatieve bedrijven, in tegenstelling tot andere toepassingen waarbij wederom gebruik wordt gemaakt van vragenlijsten zoals CIS-gegevens, maar waarbij wél de kleinere bedrijven zijn ondervraagd.

Wat de indexrapporten over het algemeen gemeen hebben is dat ze verschillende onderzoeksmethoden combineren. Dit wijst erop dat de meest complete techniek voor het identificeren van het innovatiepotentieel, inclusief de kleinere innovatieve bedrijven, een mix is van verschillende kwantitatieve en kwalitatieve methoden. Zo combineren Gauthier, et al. (2018) verschillende kwantitatieve databases die zijn samengesteld door private organisaties. Deze data wordt aangevuld met gegevens afkomstig uit vragenlijsten aan bedrijven, literatuur- of media-analyses en interviews met lokale partners of experts zoals accelerators en incubators. Op deze manier kan een relatief grote hoeveelheid aan kwantitatieve gegevens worden verzameld en gecontroleerd met behulp van kwalitatieve methoden. Er kan bijvoorbeeld een inschatting worden gemaakt hoe accuraat de verzamelde kwantitatieve gegevens zijn en voor ontbrekende gegevens kan worden bijgeschat. Nadeel van deze methode is natuurlijk de lage efficiëntie vanwege de tijdrovende toepassing van verschillende onderzoeksmethoden uitvoeren en combineren.

De meeste onderzoekstechnieken die worden gecombineerd zijn al min of meer besproken in de voorgaande paragrafen, of ze zijn slechts gericht op de resultaten en daarom minder relevant voor het meten van de input in de vorm van innovatieve activiteiten. Daarentegen bieden enkele van de big-data toepassingen door particuliere bedrijven⁶ wel interessante aanknopingspunten met betrekking tot een methode om ook de kleine bedrijven in beeld te krijgen. Deze big-data methoden zijn

Tabel 2.2 *Rapporten die innovatie potentieel indexeren.*

Indexrapport	Belangrijkste auteur(s)
Global Startup Ecosystem Report 2018	Gauthier, Stangler, Penzel, & Morelix (2018)
Global Competitiveness Index 2017-2018	Schwab (2017)
Regional Innovation Scoreboard 2016	Hollanders, Es-Sadki, & Kaner (2016)
EU Regional Competitiveness Index 2016	Annoni, Dijkstra, & Gargano (2017)
The Kauffman Index: Startup Activity	Fairlie, et al. (2015)

⁴ Opvallend is dat in de grijze literatuur wel relatief nauwkeurige tellingen te vinden zijn van innovatieve startups. Bijvoorbeeld EBU Research (2016) in Utrecht en Noseleit, de Lange, & Rambharos (2016) in Noord-Nederland. Ook houden verschillende organisaties zoals overheden, accelerators en incubators (StartupDelta, 2018) zich bezig met het opsporen van innovatie, en op het moment van schrijven heeft het Centraal Bureau voor de Statistiek [CBS] (2018) de eerste resultaten gedeeld van een big-data experiment om ook de kleine kleinere innovatieve bedrijven in beeld te krijgen.

⁵ Bijvoorbeeld het CBS en Algemeen Bedrijfsregister (ABR) in Nederland of het 'U.S. Bureau of the Census' en het 'Business Register (of SSEL)' in de VS.

⁶ Enkele voorbeelden van deze bedrijven zijn: Orb Intelligence (2018), CrunchBase (2018) en Dealroom (2018).

nog niet aan bod gekomen, en worden daarom in de volgende paragraaf nader besproken.

2.1.5 Big data

Bij big data toepassingen (zie kader 2.1 over verwarring rondom het gebruik van big data-terminologie) zoals benoemd door de auteurs van de eerdergenoemde rapporten die innovatie indexeren (tabel 2.2), is het niet altijd even inzichtelijk hoe de databases exact zijn opgebouwd. Wel wordt aangegeven dat gebruik wordt gemaakt van slimme zoekopdrachten en algoritmes die bestaande data – zoals websites, sociale media, en andere digitale bronnen – doorzoeken en vervolgens koppelen aan de juiste bedrijven. Een dergelijke methode is meestal gebaseerd op tekstuele inhoud van digitale bronnen. Deze methode van tekstanalyse heeft zijn academische wortels in *information retrieval* en de computationele taalkunde.

Bij *information retrieval* vormen documentre-

presentatie (zoals een lijst met websites) en *query*-verwerking (een *query* is een verzameling trefwoorden) de basis voor het ontwikkelen van een zoekmodel (Manning, Raghavan, & Schütze, 2009). Volgens een eenvoudig model wordt alle tekstinhoud na verwijdering van veel voorkomende stopwoorden gereduceerd tot een verzameling termen met bijbehorende absolute frequenties om de relevantie van de tekstinhoud te bepalen. Termen die onderscheidend zijn moeten echter meer gewicht krijgen en kunnen door toepassing van wegingsfuncties worden gewogen, met als resultaat een relevantiescore (Stanford Encyclopedia of Philosophy, 2014). Deze relevantiescore voorspelt dus in hoeverre de geanalyseerde tekstinhoud relevant is in verhouding tot de zoekopdracht.

De computationele taalkunde kan gezien worden als een aanvulling op het voorgaande, en classificeert woorden niet alleen op basis van hun bete-

Kader 2.1 Big data terminologie

Ondanks het erkende bestaan van veel grotere en complexere big data-toepassingen dan die in deze paragraaf worden behandeld, is er in de literatuur geen consensus bereikt over de precieze afbakening van big data. Mede door de snelle opkomst van het discours over big data, wordt de term te pas en te onpas gebruikt. Zo bestaat er veel verwarring over de fundamentele vraag hoe groot en complex de gegevens moeten zijn om te kwalificeren als 'big data' (Ward & Barker, 2013; Gandomi & Haider, 2015). De drie V's (*Volume*, *Variety* en *Velocity*) zijn tegenwoordig gebruikelijk om big data te beschrijven (Chen et al., 2012, Kwon et al., 2014), maar elke V gaat gepaard met problemen die betrekking hebben op de exacte grenzen van de kwalificatie.

In de context van deze studie blijft het gebruik van de term big data echter van toepassing en gerechtvaardigd, omdat voor elk van de 3 V's doorgaans veel computer- en servercapaciteit vereist is. Het *Volume* van de ruwe data bijvoorbeeld, is ondanks het feit dat het vaak slechts om tekstinhoud¹ gaat, groter dan veel traditionele data.² Onder de noemer *Variety* gaat het bovendien om niet-gestructureerd data die door de computer ontrokken, verwerkt en gestructureerd moet worden. Ten slotte valt er voor de derde V van *Velocity* ook het een en ander te zeggen. Afhankelijk van het aantal bijwerking van digitale bronnen of domeinnaam wijzingen, is er ook nog een mogelijkheid de veranderingen van deze data bij te houden en verwerken tot bruikbare informatie met betrekking tot de dynamiek van de gegevens.

Het gebruik van de term big data verwijst binnen de context van deze studie naar de opslag en analyse van relatief grote en of complexe gegevenssets, met behulp van een reeks technieken. In deze paragraaf gaat het bijvoorbeeld om technieken die onder de noemers *information retrieval* en computationele taalkunde geïntroduceerd worden. Er kan geconcludeerd worden dat big data een relatief concept is waarbij de classificatie – ongeacht de feitelijke grootte van de gegevens – ook grotendeels afhankelijk is van de mogelijkheden en hulpmiddelen van de gebruiker(s).

¹ Video's of afbeelding vereisen bijvoorbeeld veel grotere opslagcapaciteiten en complexere technieken om het te verwerken tot bruikbare informatie.

² Technisch gezien kunnen media- en literatuurgegevens – zoals besproken in paragraaf 2.1.3 – ook worden gezien als grote hoeveelheid data, maar een veel minder grote hoeveelheid hiervan kan met traditionele methoden worden verwerkt. Het elektronisch opslaan en verwerken van zeer grote hoeveelheden LBIO-gegevens (alhoewel het digitaliseren mogelijk veel handwerk betekent) zou dan ook gezien kunnen worden als big data toepassing.

kenissen, maar ook op basis van hun *cooccurrence*⁷ zodat ontrecte resultaten van bijvoorbeeld te algemene zoektermen zo min mogelijk worden meegenomen. In theorie kunnen zoekmodellen op basis van *information retrieval* en de computationele taalkunde net zo ver worden doorontwikkeld tot de gewenste balans tussen complexiteit en nauwkeurigheid is gevonden.

Chen, Chiang, & Storey (2012) en Kwon, Lee, & Shin, 2014 onderkennen dat big data-analyse in de afgelopen decennia steeds belangrijker is geworden binnen zowel de wetenschap als het bedrijfsleven. Wetenschappelijke artikelen die big data gebruiken om economisch- of sociaal-ruimtelijke verschijnsel en te bestuderen zijn opkomend (Kitchin, 2013; Boeing & Waddell, 2016; Zook, Shelton, & Poorthuis, 2017). Meer specifiek zijn onder de noemers *webscraping*, *web mining*, *text mining*, *data mining*, *topic modelling* en *website data*, enkele artikelen te vinden die *information retrieval* toepassen als nieuwe aanpak om grote hoeveelheden ruimtelijke gegevens van innovatieve bedrijven te genereren (Gök, Waterworth, & Shapira, 2015; Li, Arora, Youtie, & Shapira, 2016; Kinne & Resch, 2018). Het gaat om een selecte groep artikelen die tot op heden vooral het Verenigd Koninkrijk, Duitsland en de belangrijkste steden in deze landen bestuderen. De nieuwigheid en het geringe aantal wetenschappelijke studies, toont aan dat het om een nog niet wijdverbreide methode gaat voor het bestuderen van innovatie binnen de economische geografie. Eerste bevindingen van deze studies suggereren dat de methode van *webscraping* een goede aanvulling is op bestaande methoden, mits zorgvuldig uitgewerkt. Aangezien de methode van *webscraping* aansluit op de uitgangspunten van dit onderzoek, zal deze methode uitgebreid worden besproken in de volgende paragrafen.

2.1.6 Webscraping: pioniers in de geografie

De studie van Gök, et al. (2015) is de eerst gevonden wetenschappelijke poging die de geschiktheid van bedrijfswebsites onderzoekt als bron voor innovatieonderzoek. Ondanks de potentie die eerdere studies⁸ rapporteren, blijkt de literatuur over het gebruik en de geldigheid van *webscraping* relatief onderontwikkeld in vergelijking tot andere

onderzoekgebieden die *webscraping* toepassen. Gök, et al. (2015) namen daarom het initiatief om gegevens afkomstig van bedrijfswebsites te analyseren en af te zetten tegenover traditionele R&D-gegevens.

Het onderzoek van Gök, et al. (2015) biedt om te beginnen nieuwe mogelijkheden voor het identificeren en classificeren van innovatie op thema (zie ook kader 2.2). Aangezien innovatie per definitie niet binnen de traditionele sector hoeft te ontstaan, zijn traditionele sectorale indelingen (zoals de Standaard Bedrijfsindeling 2008; hierna: SBI-code) minder geschikt. De studie van Gök, et al. maakt daarom gebruik van een nieuw ontwikkeld classificatiemodel. Zij passen dit model toe op de 'groene industrie'. In theorie zou elk willekeurig thema op een vergelijkbare manier betrokken kunnen worden. De classificatie voor de 'groene industrie' is afkomstig van een eerdere studie door Shapira, Gök, Klochikhin, & Sensier (2014) en gebeurt op basis van een zoektermbenadering. De zoektermbenadering houdt in dat op basis van bestaande literatuur zoektermen (classificatie termen) worden samengesteld, waarmee bijpassende bedrijfsinformatie in een database opgespoord kan worden. Het gebruik van zoektermen leent zich vanzelfsprekend goed voor de toepassing van *webscraping* en *information retrieval*.

Aangezien bedrijven verschillen in de manier waarop zij hun bedrijfsactiviteiten en producten beschrijven, is door de onderzoekers gekozen om ruime zoektermen te gebruiken en vervolgens de bedrijven die niet goed passen weer uit te sluiten. Voordeel van de zoektermbenadering is

Kader 2.2 Geautomatiseerde classificatie

Een Koreaans onderzoek door Jung (2008), claimt dat het waarschijnlijk de eerste serieuze poging ter wereld is dat gebruik maakt van kunstmatige intelligentiemethoden voor geautomatiseerde classificatie van industrieën. De methode is vergelijkbaar met die van Gök, et al. (2015). Het CBS heeft later ook geëxperimenteerd met het automatisch classificeren van de topsectoren. De precisie (80%) van de best presterende methode was voldoende, maar nog niet nauwkeurig genoeg bevonden. Bovendien waren de subsectoren duidelijk minder goed te voorspellen dan de topsectoren (Roelands, van Delden, & Windmeijer, 2017).

⁷ *Cooccurrence* verwijst naar de bovengemiddelde kans dat twee termen beide voorkomen in een bepaalde volgorde.

⁸ Twee studies zijn noemenswaardig t.b.v. onderzoek naar innovatieve bedrijven door middel van webscraping: Youtie, Hicks, Shapira, & Horsley (2012) en Arora en Youtie, Shapira, Gao, & Ma (2013).

dat de methode kan worden aangepast op basis van de nieuwste ontwikkelingen. Nadeel is dat er maatwerk nodig is voor het samenstellen van de zoektermen, laat staan wanneer dit voor alle innovatieve sectoren moet gebeuren. En zelfs wanneer maatwerk wordt geleverd, blijft het resultaat afhankelijk van hoe goed de bedrijfsactiviteiten in de database (bijvoorbeeld de bedrijfswebsites) beschreven staan.

Naast de mogelijkheden om innovatie te identificeren op thema, wordt door Gök, et al. voornamelijk gekeken naar R&D-gegevens die afgeleid zijn van de bedrijfswebsites om vervolgens de vergelijking te maken met traditionele R&D gegevens.⁹ Om dit te doen, zijn *web-based* variabelen geoperationaliseerd om R&D-activiteiten van bedrijven te identificeren. Het onderzoek laat in eerste instantie zien dat de geïdentificeerde *web-based* R&D-gegevens significant verschillen van traditionele R&D-gegevens. Gegeven verklaringen hiervoor zijn dat websitegegevens meer gericht zijn op *downstream*-¹⁰ of klantgerichte aspecten van R&D-processen, aangezien bedrijven hun nieuwe R&D-projecten en -activiteiten pas na voltooiing op hun websites bekendmaken.¹¹

Websitegegevens zijn voornamelijk gericht op de output van R&D, terwijl traditionele R&D gegevens juist ook de input (activiteiten en investeringen) omvatten. Niet alle input leidt echter tot output, en dus zou voorspelt kunnen worden dat websitegegevens (vooral output) minder vaak R&D rapporteren dan traditionele gegevens (vooral input). Het onderzoek wijst op het tegenovergestelde. Verklaring hiervoor is dat financiële gegevens jaarlijks geaggregeerd zijn, in tegenstelling tot websitegegevens die afzonderlijke individuele R&D-activiteiten vermelden waardoor het aantal hoger uitvalt. Bovendien wordt gesuggereerd dat het verschil zou kunnen duiden op het relatief grote belang van niet-financiële input voor het R&D-proces. Deze niet-financiële input wordt wel gevangen door de gegevens op de bedrijfswebsites, maar niet door traditionele financiële R&D-gegevens. Daarnaast kunnen bedrijven er voor kiezen om niet alle details

van hun R&D-activiteiten op openbare websites te plaatsen, maar ze kunnen wel een algemene beschrijving geven van wat ze doen ten behoeve van (potentiële) klanten, zonder uitgebreide technische details aan de concurrenten te onthullen.

Gök, et al. concluderen dat websitegegevens veelbelovend zijn, juist ook voor onderzoek naar kleine innovatieve bedrijven. Uit eerder onderzoek is al gebleken dat kleinere bedrijven meestal kleinere websites hebben, waardoor het *web mining*-proces en de daaropvolgende analyse beter beheersbaar zijn (Youtie, Hicks, Shapira, & Horsley, 2012). Websitegegevens zijn ook een uitkomst voor de problemen omtrent de dekking, veroudering en respons van traditionele data waardoor doorgaans vooral de kleinere bedrijven moeilijk in kaart zijn te brengen. Bovendien zijn websitegegevens tot op zekere hoogte altijd actueel en daarom goed voor het begrijpen van bedrijfsinnovatie in snel veranderende domeinen.¹²

Maar er worden door Gök, et al. ook moeilijkheden benoemt; het interpreteren van bedrijfswebsites en de verwerking van grote datasets van websites gaat gepaard met een nieuwe reeks aan op te lossen problemen. Voor de toepassing van *web-scraping* zijn specifieke technische vaardigheden vereist die verschillen van gevestigde methoden voor innovatieonderzoek. Verder is het van belang dat zoekstrategieën zorgvuldig worden samengesteld en uitgewerkt, en dient er altijd rekening te worden gehouden met het feit dat bedrijven kunnen variëren in hun motivatie voor het plaatsen (of niet plaatsen) van informatie over hun activiteiten op websites. Zo blijkt uit een eerdere studie dat bedrijven in verschillende sectoren en landen verschillend omgaan met wat ze rapporteren, en kunnen er strategische momenten zijn waarop bedrijven meer of minder informatie publiceren over hun technologieontwikkeling of bedrijfsstrategie (Arora, Youtie, Shapira, Gao, & Ma, 2013). Toch wordt *webscraping* door alle geïnventariseerde studies als een belangrijke en nuttige aanvulling op de huidige methoden gezien. Sterker nog, het bied soms nieuwe inzichten die niet of minder snel worden verkregen met behulp van traditionele methoden.

⁹ Zoals R&D-uitgaven en -activiteiten (gerapporteerd naar aanleiding van R&D-vragenlijsten zoals CIS-enquêtes) en andere bedrijfsgegevens waaronder financiële administratie en patenten.

¹⁰ *Downstream* wordt gebruikt om iets te beschrijven dat later in het innovatieproces plaats vindt.

¹¹ Dus voor specifiek R&D-*scraping* geldt juist dat vooral resultaat (output) wordt gemeten, maar dit hangt dus van de zoektermen af.

¹² Zie het onderzoek van Blazquez & Domenech (2017) waaruit blijkt dat bedrijfswebsites een goede graadmeter kunnen zijn voor de daadwerkelijke activiteiten-status van bedrijven. Het onderzoek is uitgevoerd binnen de Spaanse dienstsector, bedrijven in andere sectoren of landen zouden afwijkende resultaten kunnen opleveren.

2.1.7 Webscraping: positieve eerste resultaten

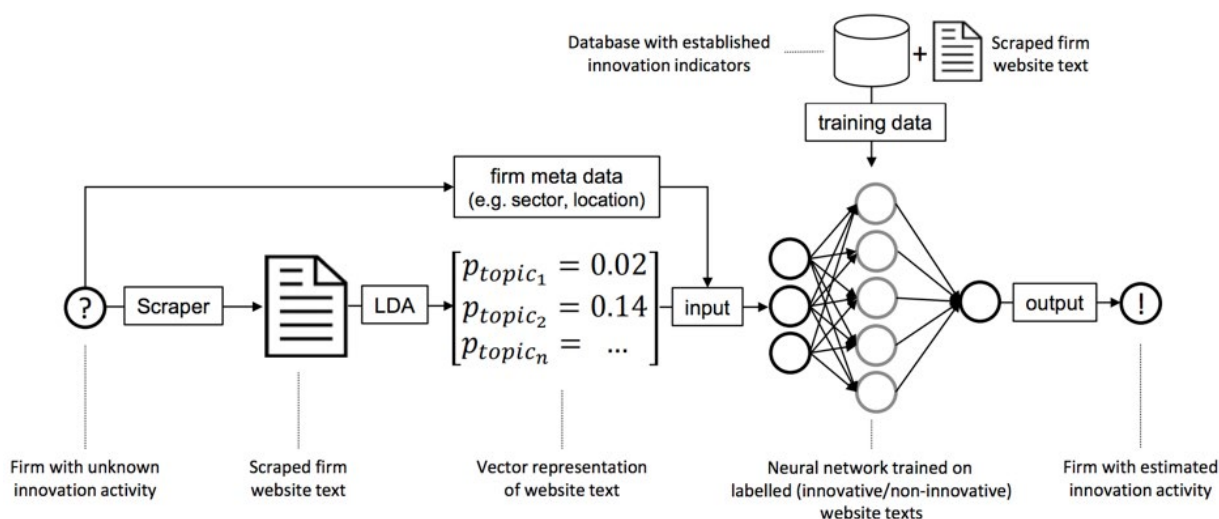
De *webscraping* methode zoals deze door Gök, et al. is onderzocht, heeft vooralsnog slechts navolging gekregen in de vorm van enkele experimentele en verkennende studies. Deze studies gebruiken *webscraping* voor onderzoek naar innovatie bij bedrijven, maar richten zich vaak op een specifieke industrie (Möller, 2014; Beaudry, Héroux-Vaillancourt, & Rietsch, 2016; Li, et al., 2016; Shapira, Gök, & Salehi, 2016; Héroux-Vaillancourt & Beaudry, 2018) of op aspecten van de *webscraping* methode (Mennis & Guo, 2009; Blazquez & Domenech, 2018). De resultaten van deze studies bevestigen opnieuw dat *webscraping* nieuwe en aanvullende inzichten oplevert, maar ook dat deze nog in ontwikkeling is. Dit laatste is een mogelijke verklaring voor het feit dat het potentieel nog niet volledig wordt benut. De enige gevonden studies die daadwerkelijk grote hoeveelheden ruimtelijke gegevens van alle innovatieve bedrijven in een regio genereren door middel van *webscraping* staan nog in de kinderschoenen en zijn pas in de zomer van 2018 in eerste vormen gepubliceerd. Het gaat om tussenresultaten van twee afzonderlijke projecten die als einddoel hebben om een regionaal overzicht te genereren van (kleine) bedrijven die actief zijn op verschillende innovatieve gebieden. Deze voorbeelden zullen hieronder worden besproken.

Bij het eerste project, *scrapen* Kinne & Resch (2018) ruim 2,3 miljoen websites van Duitse bedrijven die ingeschreven staan in het handelsregister. Vervolgens wordt een inschatting gemaakt hoe innovatief de bedrijfsactiviteiten zijn op basis van de teksten op de bedrijfswebsite (zie figuur 2.2). Dit doen ze door eerst een deel van de onderzoekspro-

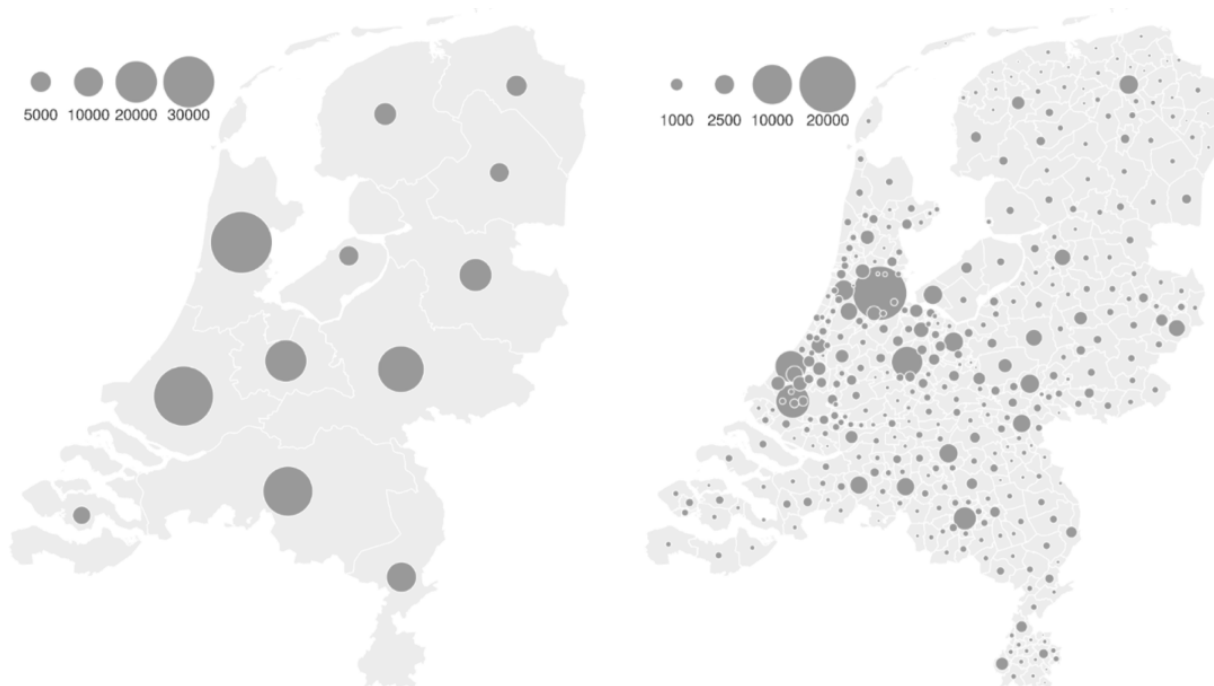
pulatie – waarvan al andere verzamelde informatie beschikbaar is over de innovatieve activiteiten – te *scrapen* en vervolgens de tekstinhoud te analyseren met behulp van topic-modelling (*Latent Dirichlet Allocation* of LDA) en te trainen op basis van de *Artificial Neural Network* (ANN) methode.¹³ De getrainde algoritmes zijn vervolgens gebruikt om het totaal aan beschikbare websites te analyseren, met als voorlopig resultaat een database van alle websiteteksten en een ranglijst van innovativiteit. Het is slechts een eerste stap binnen een groot-schalig langetermijnonderzoek, enkele technische problemen moeten nog worden verholpen en aan de verbetering van algoritmes en een rigoureuze kwantitatieve validatieprocedure moet nog worden gewerkt. Als deze stappen voltooid zijn kan de verspreiding van innovatieve en niet-innovatieve bedrijven op een gedetailleerd ruimtelijk schaalniveau nader bestudeerd worden. Voor nu, zijn vooral de methodologische inzichten van waarde voor de opzet en uitvoering van dit onderzoek.

Het tweede project omvat ook slechts eerste resultaten van een nog te verfijnen onderzoek, ditmaal naar buiten gebracht als nieuwsbericht door het CBS (2018). Hierin wordt verkondigd dat het CBS sinds kort experimenteert met het gebruik van big data in de vorm van *webscraping* (zie figuur 2.3 voor de overzichtskaarten van het resultaat). Zij doen dit vooral om de kleine bedrijven – waarvan eerder al is geconstateerd dat veel informatie ontbreekt – in beeld te krijgen. Het feit dat dit nu pas, en gelijktijdig met dit onderzoek gebeurt, onderstreept de behoefte om de onderzoekskloof

¹³ Voor een uitgebreide beschrijving, zie de methodebeschrijving in Kinne & Resch (2018, p. 85-87)



Figuur 2.2 Voorgesteld model om de innovatieactiviteit van bedrijven te schatten op basis van hun websiteteksten. Bron: (Kinne & Resch, 2018)



Figuur 2.3 Aantal kleine innovatieve bedrijven per provincie (link) en gemeente (rechts). Bron: (CBS, 2018)*

* Inmiddels werkt de Innovatiespotter samen met de mensen achter dit onderzoek (samenwerkingspartners in een onderzoek van Economische Zaken). Uit interne informatie blijkt dat de Innovatiespotter in vergelijking tot de gehanteerde methode door het CBS (2018), m.b.t. innovativiteit andere eisen stelt aan de zoekopdracht waardoor het tot minder resultaten komt (G. Vellinga, persoonlijke communicatie, 17 februari 2019). Dit bevestigt de kwalitatieve aard van innovatie die leidt tot onvergelijkbare kwantificeringen.

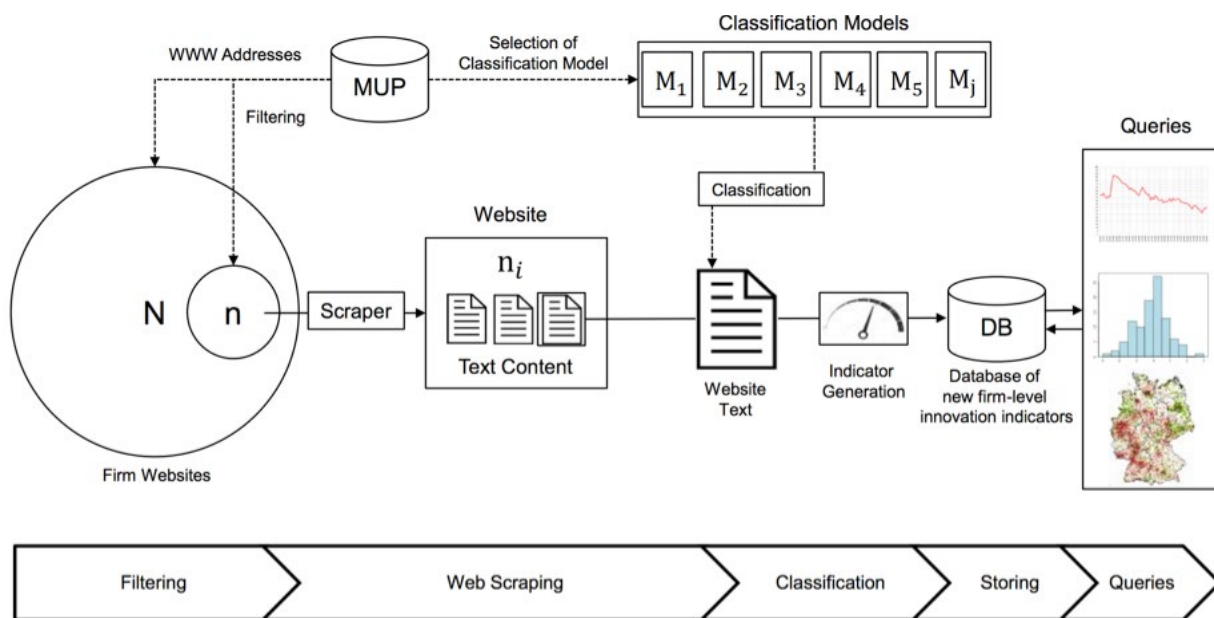
tussen de grote en kleine bedrijven te dichten. In het CBS onderzoek zijn een half miljoen bedrijven met minder dan 10 werkzame personen uit het bedrijvenregister geanalyseerd. Evenals het onderzoek van Kinne & Resch, is het algoritme getraind op basis van bestaande innovatie gegevens van (grote) bedrijven. De werkwijze van zowel het CBS, Kinne & Resch als externe private partijen (voor zover zij inzicht verschaffen in hun werkwijze) vertonen veel overeenkomsten met het werk van Gök, et al. (2015). De methodologische inzichten en resultaten van al deze studies kunnen worden meegenomen in de opzet en analyse bij het vervolg van dit onderzoek.

Kortgezegd is het grote voordeel van de *web-scraping* methode dat grote hoeveelheden data goed toegankelijk zijn en relatief eenvoudig door een computermiddel kunnen worden verwerkt. Grootste nadeel is dat niet van elk bedrijf informatie makkelijk online te vinden is, en dat ook minder relevante bedrijven gevonden kunnen worden. Doorgaans wordt dit laatste probleem opgelost door (een deel van) de automatisch behandelde website teksten handmatig te checken. Dit kan door zelf de website informatie kritisch te bekijken met of zonder hulp van aanvullende controlegegevens. Ook kan het bedrijf zelf worden benaderd om naar de juistheid en volledigheid van de website infor-

matie te informeren. Vervolgens kunnen steeds betere zoektechnieken en algoritmes het menselijke werk overnemen. In theorie kan dit steeds verder doorgevoerd worden, maar de ontwikkeling en uitvoering zal nog steeds veel tijd in beslag nemen; zeker wanneer alles up to date gehouden moet worden. Op een gegeven moment kan een grens bereikt worden waarbij de algoritmes dusdanig ingewikkeld worden dat het praktisch onwerkbaar wordt of te veel problemen met zich meebrengt.

2.1.8 Webscraping: toepassing per thema

Tot dusver zijn verschillende methoden behandeld die gebruikt worden om de innovativiteit van bedrijven te meten. Daarbij is vooral focus gelegd op de innovatie input in de vorm van activiteiten. Het blijft echter vaag hoe innovatief de daadwerkelijke activiteiten precies zijn, het doel is dit zo goed mogelijk in te schatten. Met *webscraping* kan dit op twee manieren worden aangepakt. De eerste manier is classificatietermen samenstellen die duiden op innovatieve activiteiten ongeacht de sector, en is in de voorgaande paragrafen besproken. Bij de tweede manier worden classificatietermen samengesteld om juist de activiteiten te verbinden aan een sector. Vallen bedrijfsactiviteiten binnen een innovatieve sector, dan wordt het bedrijf automatisch bestempeld als innovatief.



Figuur 2.4 *Procesbeschrijving: webscraping, tekstclassificatie, genereren van innovatie-indicatoren, opslaan en opvragen van data.* Bron: (Kinne & Resch, 2018)

In het onderzoek van Kinne & Resch (2018) wordt zowel de mate van innovativiteit als de sector geanalyseerd (figuur 2.4). Volgens het CBS (2018) zijn sommige sectoren makkelijker te voorspellen op basis van de websitegegevens dan andere; zo zouden vooral technologisch innovatieve bedrijven goed te voorspellen zijn. Het identificeren van innovatie op basis van de sector begint bij het kiezen van een geschikte bedrijfsindeling. Er bestaan verschillende classificatiemethoden voor de categorisering van bedrijfssectoren. De tot dusver behandelde literatuur geeft geen uitsluitsel over de beste of meest gangbare manier van classificeren. Dit lijkt per land te verschillen en is ook afhankelijk van het onderliggende doel van de afzonderlijke onderzoeken.

Al eerder is vastgesteld dat voor innovatie, de traditionele sectorale indelingen (zoals SBI-industrie code in Nederland) minder geschikt zijn omdat ze niet altijd rekening houden met de nieuwste trends rondom innovatie.¹⁴ Het ligt daarom meer voor de hand een aangepaste sectorale indeling te gebruiken voor de identificatie en analyse van innovatie. Gauthier, et al. (2018) van Startup Genome erkennen dat sommige startups zich bevinden in subsectoren die niet meegenomen zijn in hun analyse. Ook het feit dat de door Startup Genome gehanteerde subsectoren elkaar niet wederzijds uitsluiten, bevestigt de moeilijkheid om een duidelijk onderscheid aan te houden dat ook nog eens eenduidig is.

¹⁴ Indeling wordt elke 15 tot 20 jaar aangepast (CBS, 2017).

In Nederland wordt regelmatig de topsectoren indeling toegepast als het gaat om innovatie classificeren. Deze topsectoren werden in 2011 geselecteerd via een bottom-up-proces waarin publieke en private actoren zich konden presenteren als kandidaten (CBS, 2014). Het CBS bakent op basis van de SBI-codes deze topsectoren af. Bij deze indeling sluiten de categorieën elkaar eveneens niet uit aangezien sommige SBI-codes in meerdere topsectoren kunnen voorkomen. Hoewel de SBI een vrij gedetailleerde classificatie is, verricht het CBS voor een aantal sectoren aanvullend maatwerk omdat in sommige gevallen de SBI-classes toch te breed zijn (CBS, 2017). Daarnaast heeft het gebruik van de SBI-code als tekortkoming dat slechts één hoofdactiviteit bekend is. Bedrijven, en met name de innovatieve, kunnen echter actief zijn in verschillende sectoren en zij kunnen ook in (hoofd) activiteit(en) variëren. Kortom, de SBI-code klopt niet altijd. Zoals eerder aangekaart, kan met *webscraping* een alternatieve indeling gemaakt worden die wel in staat is om recentere en eventuele aanvullende innovatieve activiteiten die het bedrijf kenmerken mee te nemen. Dit kan op basis van een zoekterm benadering zoals in het eerder besproken onderzoek van Gök, et al. (2015).

Dit eerste deelhoofdstuk heeft laten zien dat verschillende methoden gangbaar zijn en elkaar aanvullen als het gaat om innovatie bij bedrijven opsporen. Beperkingen die verband houden met veelgebruikte methoden voor innovatieonderzoek hebben relatief vaak of in grotere mate betrekking op kleinere bedrijven. Het is belangrijk rekening te

houden met deze beperkingen, niet alleen in de voortzetting van dit onderzoek, maar ook voor het interpreteren van bestaand onderzoek. In vergelijking tot andere methoden lijkt *webscraping* veelbe-

lovend om verschillende innovatieve activiteiten van de kleine en nieuwe bedrijven in een vroeg stadium op te sporen. Dit is doorgaans moeilijk of zelfs onmogelijk met de meer gebruikelijke methoden.

2.2 Meten van Startups

Het meten van startups kan in twee stappen worden uitgelegd die in dit deelhoofdstuk achtereenvolgens besproken zullen worden. Als eerste stap moet worden bepaald wat er precies wordt gemeten. Op basis van de literatuur zullen verschillende startupdefinities worden bestudeerd, inclusief de consequenties die verbonden zijn aan de afzonderlijke keuzemogelijkheden. Elke keuze wordt beïnvloed door de voor- en nadelen van bijpassende methoden en heeft consequenties voor stap twee: de interpretatie. Voor deze tweede stap wordt ter afsluiting van dit deelhoofdstuk besproken hoe er in het kader van daadwerkelijk innovatie meten een betekenis kan worden gegeven aan tellingen van startup bedrijven.

2.2.1 Startup definities

Tegenwoordig staat de term 'startup' internationaal bekend als motor voor economische groei, maar er bestaat geen academische consensus over de exacte betekenis van startups en de reikwijdte van de dekkingsgrens: "*A review of literatures shows that 'startup' is a subset of 'entrepreneurship', and the important issue of 'startup' is innovation-related inventions in various aspects.*" (Baimai, 2017, p. 10) Een mogelijke verklaring voor het ontbreken van duidelijke definities in de academische literatuur, is dat het concept vooral veel in het bedrijfsleven en de volksmond wordt gebruikt. In de wetenschap lijkt men te kiezen voor '*new firms*', met eventueel toevoegingen – zoals '*entrepreneurial*', '*smaller*' of '*high growth*' – afhankelijk van waar de nadruk op ligt.

Er zijn verschillende manieren om startup bedrijven te kwalificeren en kwantificeren. Volgens Baimai (2017) biedt het bedrijfsleven een duidelijkere en meer omvattende definitie van de betekenis van het concept startup. De internationaal bekende definitie van Steve Blank (2010) definieert een startup als een bedrijf, samenwerkingsverband of tijdelijke organisatie die is ontworpen om te zoeken naar een herhaalbaar en schaalbaar businessmodel (Spender, Corvello, Grimaldi, & Rippa, 2017). Onder andere Gauthier, et al. (2018) maakt gebruik van deze definitie.

Dat met een startup een nieuw bedrijf of organisatie bedoeld wordt is onomstreden. Voordat een meting kan worden gestart, moet echter een afgebakende wetenschappelijke definitie worden gekozen die niet alleen uitsluitel geeft over hoe het aspect 'nieuw' kan worden geoperationaliseerd, maar ook duidelijk maken welke aanvullende kenmerken vereist zijn om de kwalificatie startup te krijgen. Luger & Koo (2005) merken op dat een startup tot op zekere hoogte altijd nieuw, actief en onafhankelijk moet zijn. Dit komt ook terug in de vier uitdagingen die Mason (1983) beschrijft bij het definiëren van startups – of *new firms* om in zijn eigen bewoording te spreken. Deze uitdagingen hebben betrekking op aspecten van het vormingsproces: de opstartdatum, *newness* of relatieve nieuwheid, vroege veranderingen van activiteit en onafhankelijkheid. De volgende paragrafen bespreken elk van deze aspecten.

2.2.2 Opstartdatum

Wat betreft de vaststelling of een onderneming nieuw is, wordt door Mason (1983) onderscheid gemaakt in de opstartdatum (absoluut) en of een onderneming daadwerkelijk nieuwe dienst- of productie capaciteit toevoegt (relatief). De exacte opstartdatum van een onderneming is niet altijd duidelijk te bepalen. De opstart is vaak een langdurig proces. In veel gevallen opereren opkomende bedrijven eerst vanuit een informele werkplek en vaak ook in deeltijd. Alleen wanneer de levensvatbaarheid van de onderneming is aangetoond wordt de bedrijfslocatie formeel en beginnen oprichters de bedrijven fulltime te leiden. Veel startup definities zijn juist gericht op de opstartfase en de bijhorende bedrijfskenmerken van dat moment. Voor de bepaling van 'nieuw' in absolute zin, staat daarom het kiezen van een meetpunt in de vorm van de opstartdatum centraal.

De gebruikelijke benadering, is om de opstartdatum van een nieuw bedrijf gelijk te stellen met het moment van registratie in het handelsregister.¹⁵ Registratie geeft echter geen garantie dat het be-

¹⁵ In sommige gevallen worden licenties gebruikt, maar daar kleven ook weer nadelen aan zie Luger & Koo (2005, p.19).

drijf daadwerkelijk actief is. Sterker nog, sommige geregistreerde bedrijven bestaan slechts op papier om juridische, fiscale of andere administratieve redenen. Daarbovenop is het volgens Mason (1983) problematisch dat sommige ondernemingen niet ingeschreven staan terwijl ze al meerdere jaren actief handelen. In Nederland is elke onderneming echter verplicht om zich in te schrijven zodra sprake is van deelname aan het economische verkeer (Kamer van Koophandel, zd). Toch kan sprake zijn van zelfstandig handelen terwijl er op papier nog geen nieuwe inschrijving bestaat – bijvoorbeeld in de vorm van een projectgroep dat niet te herleiden is naar één specifieke inschrijving. Daarnaast zullen bedrijven die niet lang genoeg overleven om de fase van actieve economische deelname te bereiken, nooit verplicht zijn zich in te schrijven en daarom soms ontbreken in de registers. Voor veel onderzoek is het ontbreken van deze laatste groep niet problematisch aangezien veel startup definities vereisen dat de onderneming al wél betrokken is bij de handel in goederen of diensten (Luger & Koo, 2005).

Een andere methode, is om het in dienst nemen van de eerste werknemer als startdatum te nemen. Uiteraard sluit dit organisatievormen zoals de eenmanszaken waarbij geen medewerkers in loondienst zijn uit. In welke mate dit problematisch is hangt af van de gangbare werkvormen in een sector. Zo zijn in sommige sectoren relatief veel bedrijven te vinden die alleen betaald werk voor de oprichters opleveren. Zo verwijst Mason (1983) naar onderzoek waaruit blijkt dat familie van de oprichter optreedt als een bron van onbetaalde werk. Blevins & Ragozzino (2018) benoemen het vele voorkomen van samenwerkingsverbanden binnen het startup milieu. Dergelijke constructies verschaffen niet altijd inzicht of duidelijkheid over de feitelijke werknemers, zeker in gevallen waar werknemers niet in loondienst worden genomen. Voor zowel de datum van registratie in handelsregisters als het in dienst nemen van de eerste werknemer, geldt dat aanvullend maatwerk nodig is om de validiteit te controleren. Hierbij dient een balans tussen efficiëntie en nauwkeurigheid gevonden te worden die aansluit bij het uiteindelijk onderzoeksdoel en de bijhorende onderzoeksmethode(n).

2.2.3 Nieuwheid

Naast de opstartdatum die nodig is om in tijdseenheden vast te kunnen stellen wanneer een bedrijf nieuw is, heeft het woord nieuw ook nog een relatieve betekenis. Dit wordt door Mason aangeduid

als de ‘nieuwheid’ en uitgelegd als het toevoegen van nieuwe dienst- of productie capaciteit aan de economie. Hoewel een bedrijf op papier nieuw is, hoeft er niet altijd sprake te zijn van een nieuwe combinatie van middelen. Bijvoorbeeld wanneer het management van een dochteronderneming zelf eigenaar wordt om vervolgens als onafhankelijk bedrijf gescheiden van het moederbedrijf verder te gaan. Andersom kan het ook zo zijn dat situaties van opvolging of rigoureuze veranderingen binnen een bedrijf, hoewel gebaseerd op reeds bestaande juridische entiteiten, in elke andere zin als nieuwe onderneming gezien kunnen worden omdat ze een totaal nieuwe rol in de economie vervullen.

Met de startdatum als beginpunt, kan een grens worden toegevoegd tot wanneer een bedrijf als nieuw gezien wordt en wanneer dit niet meer het geval is. Veel onderzoeken en definities kijken uitsluitend naar het moment dat de bedrijven ontstaan. Hoe lang een bedrijf de kwalificatie startup mag behouden is dan niet relevant en wordt buiten beschouwing gelaten. Katila, Chen, & Piezunka (2012), bijvoorbeeld, definiëren ‘*entrepreneurial startups*’ als bedrijven die vanuit een zwakke markt en slechte toegang tot hulpmiddelen opereren. Zij meten startups uitsluitend op het moment van ontstaan. In sommige gevallen is het wel van belang om oud van nieuw te scheiden. Dit vereist vaak aanvullende informatie over de te beoordelen bedrijven. De hoeveelheid benodigde kennis, is sterk afhankelijk van de definitie die wordt aangenomen.

Zo is een startup volgens Graham (2012) een bedrijf dat opgericht is om snel te groeien. Wordt het grootste deel van de groei in de toekomst voorzien, dan blijft de kwalificatie startup van toepassing. Aansluitende definities en omschrijvingen uit de wetenschap wijzen er op, dat het bereiken van volwassenheid dient als omslagpunt. Volgens Lipi (2013) is volwassenheid bereikt wanneer sprake is van een bureaucratische organisatiestructuur, werkbeschrijving, beleid en procedures. In de volwassenheidsfase zijn hiërarchische relaties veel formeler geworden. Volwassen bedrijven hebben de overlevingstest met succes doorstaan, en proberen te beschermen wat ze hebben verkregen in plaats van zich op nieuwe terreinen te begeven. Scott & Bruce (1987) voegen hier aan toe dat een bedrijf zijn opstartfase is ontgroeit zodra sprake is van een afhankelijk management, het bedrijf eigendom is geworden van een grotere groep mensen of wanneer het lokale werkgebied van het bedrijf is uitgedroeit tot een uitgebreider werkgebied. Er

kan kortom worden gesteld dat er sprake is van een spectrum waarin in meer of mindere mate sprake kan zijn van een startup. Dit suggereert dat het misschien niet logisch is om binnen dit spectrum een al te harde grens aan te houden.

Ten behoeve van onderzoek kan het echter wel nuttig zijn om een harde grens te trekken tussen een startup en een volwassen bedrijf. Het is voor de inventarisatie van grote aantallen bedrijven niet praktisch om alle bovengenoemde kenmerken van volwassenheid per bedrijf te identificeren. Vanuit het bedrijfsleven en de grijze literatuur bestaan ook indicatoren voor volwassenheid die makkelijk te achterhalen zijn. Bijvoorbeeld het niet langer genereren van risicokapitaal en het bereiken van de beursgang. Ook kunnen praktische indicatoren, die volgens veel definities weliswaar niet bepalend zijn voor het bereiken van volwassenheid, toch bruikbaar zijn omdat ze beter toe te passen zijn voor grotere onderzoeken en wel goede voorspellers kunnen zijn van dat wat moet worden gemeten. Voorbeelden van dit soort voorspellers zijn het aantal werknemers, de omzet of de leeftijd van het bedrijf.¹⁶ De algemene veronderstelling is dat hoe groter en ouder het bedrijf is, hoe minder waarschijnlijk dat er nog steeds sprake is van een startend bedrijf.

Samengevat, geldt dat een betrouwbaar, herhaalbaar, schaalbaar en winstgevend bedrijfsmodel – ongeacht de grootte (werkzame personen) en omzet – volgens de literatuur gelijk staat aan het bereiken van volwassenheid. Een volwassen bedrijf is geen startup meer, en andersom geldt dat onvolwassen kenmerken, er op wijzen dat het wél om een startup gaat. Dit laatste kan verwarrend zijn wanneer een nieuw bedrijf veel kenmerken van volwassenheid vertoont. In zo'n geval is de exacte definitie die wordt aangehouden bepalend, en het mag duidelijk zijn dat hier enige variatie en keuzevrijheid in bestaat. Daadwerkelijk vast stellen in welke mate sprake is van een 'onvolwassen' en dus nieuw bedrijf kan een complexe en tijdrovende taak zijn. Voor de analyse van grote hoeveelheden bedrijven ligt het daarom meer voor de hand om een definitie te kiezen of indicatoren te gebruiken die een efficiënte meting mogelijk maken van wat nieuw is, ook al gaat het kiezen voor efficiëntie ten

¹⁶ Onderzoek van Stel & Suddle (2008) concludeert dat het maximale effect van nieuwe bedrijven op regionale ontwikkeling na ongeveer zes jaar wordt bereikt. EBU Research (2016) stelt bij de operationalisering dat een startup maximaal 20 werknemers mag hebben en maximaal 5 jaar oud mag zijn.

koste van de nauwkeurigheid.

2.2.4 Dynamische activiteiten

Een volgende uitdaging die door Mason benoemt wordt, betreft veranderingen van bedrijfsactiviteiten door de tijd. Mason (1982) suggereert dat dit juist bij nieuwe bedrijven relatief vaak voorkomt. Zo zouden bedrijven activiteiten kunnen wijzigen of misschien zelfs overgaan tot geheel nieuwe activiteiten, activiteiten uitvoeren die voorheen uitbesteed werden of juist zelf gaan uitbesteden. De problematiek rondom het meten van bedrijfsactiviteiten is in het voorgaande hoofdstuk – in het kader van de innovatieve activiteiten – al uitgebreid besproken. Om niet in herhaling te vervallen, volstaat de vermelding dat, met name in de beginjaren, extra voorzichtigheid geboden is bij het gebruik van gegevens die zijn gebaseerd op activiteiten op een bepaald moment. De identificatie en analyse van dynamische kenmerken van nieuwe bedrijven moet bij voorkeur gebaseerd zijn op meer dan alleen een meetpunt rondom de startdatum.

2.2.5 Het belang van onafhankelijkheid

Tot nu toe zijn de kenmerken 'nieuw' (opstartdatum en nieuwheid), 'actief' (deelname aan economisch verkeer) en de 'dynamische activiteiten' die een bedrijf uitvoert (vroegge veranderingen van activiteit) behandeld. Als laatste belangrijke pijler wordt 'onafhankelijkheid' besproken. Dit wordt vaak betrokken in de definitie omdat bedrijven die door gevestigde bedrijven worden opgericht aanzienlijk verschillen met bedrijven die zijn opgericht door individuen. Ze hebben te maken met verschillende economische problemen en mogelijk verschillende economische drijfveren (Johnson, 1978). Onafhankelijke nieuwe bedrijven zullen eerder vernieuwende ideeën in een regio introduceren die van invloed kunnen zijn op 'bestaande structuren' en daarmee de regionale economie op grotere schaal kunnen stimuleren (Luger & Koo, 2005). Het is daarom gebruikelijk om de categorieën afzonderlijk te identificeren door een definitie te hanteren die nadruk legt op de onafhankelijkheid van bijvoorbeeld een moederbedrijf.

Het maken van dit onderscheid in onafhankelijkheid blijkt in de praktijk echter niet altijd mogelijk. Zelfs als een strikte definitie wordt aangehouden die uitsluitend ondernemingen omvat die zijn opgericht door individuen of groepen van individuen, lost dit de onafhankelijkheidsproblematiek niet geheel op. Ook onder de kleine bedrijven is namelijk regelmatig sprake van (in)formele manage-

ment- en eigendomsverhoudingen tussen juridisch onafhankelijke bedrijven (Mason, 1982). Dit laatste neemt niet weg dat de meest formele banden relatief eenvoudig kunnen worden vastgesteld om zodoende een onderscheid aan te houden. Volgens de Schumpeteriaanse literatuur zijn onafhankelijke bedrijven beter in staat om kennis dat door gevestigde bedrijven verwaarloosd of over het hoofd wordt gezien om te zetten in maatschappelijk nut. Met andere woorden, bij een onafhankelijk bedrijf is volgens deze theorieën de kans op disruptieve innovaties veel groter.

Een nuancering die aan het voorgaande kan worden toegevoegd is dat startups die formeel afhankelijk zijn van een gevestigd bedrijf, nog steeds innovatief en zelfs disruptief kunnen zijn. Het gaat daarom te ver om alle afhankelijke startup bedrijven maar weg te laten. Sterker nog, aanvullende theorieën benoemen de kracht van 'intrapreneurs' binnen de context van grotere gevestigde bedrijven. Intrapreneurship wordt beschreven als de ondernemende actie binnen een gevestigde organisatie en richt zich voornamelijk op het opzetten van nieuwe ondernemingen (Gapp & Fisher, 2007). In tegenstelling tot de onafhankelijke ondernemer, handelt de intrapreneur binnen een bestaande organisatorische context en richt zich in essentie op het genereren van nieuwe bedrijfsgebieden (Cameki-Ordaz et al., 2012). De term afhankelijkheid met betrekking tot (disruptieve) innovatie slaat dus niet zo zeer om strikt formele onafhankelijk, maar betreft bij 'afhankelijke' startups vooral de mate waarin er sprake is of kan zijn van intrapreneurship. Volgens deze redenering is de doorslaggevende factor voor innovatie dan de mate van intrapreneurship dat binnen de startup aanwezig is of gestimuleerd wordt door het moederbedrijf. Dit kan worden omschreven als getolereerde onafhankelijkheid, maar ook als 'positieve afhankelijkheid' wanneer het moederbedrijf een ondersteunende rol aan neemt en middelen verstrekt aan de startup.

Onderscheid in formele onafhankelijk blijft zinvol. Niet omdat het een per definitie beter is dan het andere, maar omdat intrapreneuriale eigenschappen over het algemeen slechts beperkt tot uiting kunnen komen binnen de gevestigde structuren. Het vaststellen van intrapreneuriale kenmerken vergt vaak ook meer onderzoek. Een indicatie van formele (on)afhankelijkheid ten behoeve van onderzoek is daarentegen soms eenvoudig vast te stellen. Bovendien sluit het de voor innovatie

minder interessante vormen zoals fusies of filialen per definitie uit. Een indicatie van onafhankelijkheid kan daarnaast een goede voorspeller zijn van de mate waarin innovaties kansrijk zijn. Niet voor niets is een veel gehoord geluid in de literatuur dat nieuwe kansen vaak beter via het onafhankelijke ondernemerschap worden benut. Nieuwe dochterondernemingen worden mede daarom vaak niet in het onderzoek opgenomen (Stel & Suddle, 2008), afhankelijk van het doel is het echter de vraag wanneer en of dit terecht is.

Het niet opnemen van nieuw opgezette dochterondernemingen neemt namelijk niet weg dat dochterondernemingen die een afhankelijke relatie met het moederbedrijf behouden, toch innovatief kunnen zijn. Ondanks de bedreigingen die een nieuwe dochteronderneming kan oproepen voor de gevestigde structuren, is de paradox dat dit soort nieuwe formaties zich nog steeds vormen. Juist omdat ook gevestigde bedrijven zich realiseren dat innovatie nodig is en vaak een afwijking van de conventionele benaderingen vereisen (Gapp & Fisher, 2007). Studies benadrukken daarom de voordelen van het intrapreneurship. Als bestaande bedrijven een 'veilige' of vergevingsgezinde omgeving bieden waarin startup-activiteiten geen hinder ondervinden, kan dit een aanvulling zijn op de onafhankelijke startup. Zo stelt Parker (2011) dat bedrijven die een hoge mate van ondersteuning bieden aan werknemers, juist ook kansen benutten voor diegene die niet het vereiste vertrouwen hebben om onafhankelijk te ondernemen. Echter is hij ook de eerste om te verwijzen naar onderzoek van Hisrich (1990) en Hitt, et al. (2002) die de suggestie opwerpen dat intra- en entrepreneurs in dit opzicht nogal op elkaar lijken. Een studie van Mason (2013a) komt tot de conclusie dat de verschillen tussen bedrijfs-*spin-offs* en startups zeer beperkt zijn; beide diversifiëren de regionale economie door nieuwe markten te betreden. De vraag of innovaties beter door intrapreneurs of entrepreneurs tot bloei komen blijft vooralsnog onbeantwoord. In de literatuur worden wel regelmatig verschillen benadrukt, maar zowel onafhankelijke als afhankelijke startups kunnen theoretisch over dezelfde drijfveren en hulpmiddelen beschikken. Er is vooralsnog onvoldoende empirisch onderbouwd bewijs met betrekking tot de voorkeur voor het een of het ander op basis van vergelijkingen tussen intra- en entrepreneurschap.

Enkele studies trekken wel de vergelijking. Parker (2011) onderzoekt de factoren die bepalen of nieuwe zakelijke kansen eerder worden benut

door een nieuwe onderneming te starten voor een bestaande werkgever ('beginnend intrapreneurshap') of onafhankelijk van gevestigde bedrijven ('beginnend ondernemerschap'). Ook hij stelt dat er tot op heden slechts beperkt empirisch bewijs bestaat over de factoren die onafhankelijk ondernemerschap bevorderen ten gunste van intrapreneurshap. Wel concludeert hij dat er duidelijke verschillen zijn, zoals de bevinding dat onafhankelijke startups eerder geneigd zijn zich direct tot de individuele klanten te richten terwijl intrapreneuriale startups eerder geneigd zullen zijn andere bedrijven te bedienen. Uit een onderzoek naar High Tech bedrijven binnen de Nederlandse context, blijkt dat intrapreneurs en de bijhorende intrapreneuriale geest het beste tot recht komen binnen de bedrijfscultuur van kleinere organisaties (Veenker, Van der Sijde, During, & Nijhof, 2004). Dit is in lijn met de overweldigende hoeveelheid literatuur waarin wordt geconcludeerd dat nieuwe en kleine bedrijven een onevenredig grote rol spelen bij het tot stand brengen van (radicale) innovaties (Braunerhjelm, Ding, & Thulin, 2018), de status van onafhankelijk is dan mogelijk slechts van secundair belang.

Een ander onderscheid wordt gemaakt door Alvarez-Garrido & Dushnitsky (2016) en betreft het verschil in steun met onafhankelijk startkapitaal of bedrijfskapitaal. De studie gebruikt data over de innovatiegraad van startups in Amerikaanse biotechnologiebedrijven die tussen 1990 en 2003 zijn opgericht. De auteurs stellen dat onafhankelijk startkapitaal een belangrijke rol speelt, maar constateren een trend waarbij dit kapitaal in mindere mate beschikbaar wordt gesteld terwijl de interesse van bedrijven om te investeren in startups toeneemt. Dit komt overeen met uitspraken uit de literatuur over het toenemende bewustzijn bij gevestigde bedrijven over de noodzaak van innovatie om te kunnen blijven concurreren. De nadruk op basis van de studie van Alvarez-Garrido & Dushnitsky ligt op het feit dat onafhankelijke startups vaak slechts een financiële relatie onderhouden met de investeerders, terwijl startups gesteund door bedrijven in veel gevallen ook aanvullende middelen aangereikt krijgen. In het geval van de biotechnologie gaat het dan om complementaire middelen zoals een grote talentenvijver voor onderzoek en ontwikkeling (R&D), laboratorium infrastructuur en toegang tot kennis met betrekking tot regelgeving. Dit resulteert volgens de onderzoekers in meer innovatie voor startups die naast financiële steun ook toegang kunnen krijgen tot complementaire

middelen.

Ook andere auteurs verwijzen naar het belang van verschillende soorten complementaire middelen verstrekt door een gevestigd bedrijf om innovatie succesvol te kunnen ontwikkelen en op de markt te brengen. Naast financiële steun beslaan deze middelen expertise en infrastructuur voor productontwikkeling, productie, juridische zaken, verkoop, distributie en klantenservice (Colombo, Dawid, Piva, & Vivarelli, 2017). Bosma & Wenekers (2011) bevestigen dat intrapreneuriale startups in vergelijking tot onafhankelijke startups naast hogere aspiratieniveaus betere toegang hebben tot diverse middelen om groei te kunnen realiseren.

Kortom, niet strikt formele onafhankelijkheid, maar functionele onafhankelijk om vrij te kunnen innoveren blijkt doorslaggevend voor de kwalificatie van een innovatieve startup. Onderscheid in formele afhankelijkheid biedt vooral implicaties voor de mogelijke steun waar een startup toegang tot heeft en het type innovaties dat kan worden verwacht. Radicale innovaties zullen sneller geïnitieerd worden door onafhankelijke startups, maar formeel afhankelijke startups zijn onder de juiste omstandigheden ook in staat om radicale innovaties te ontwikkelen. Onderscheid in afhankelijkheid is zeker niet allesbepalend wanneer het meten van potentiële innovatie ten doel wordt gesteld. Voor niet-radical innovatie geldt misschien zelfs dat de steun van een moederbedrijf een bepalende succesfactor kan zijn voor succesvolle innovaties. Voor de onafhankelijke startup is de toegang tot financiële en complementaire middelen minder vanzelfsprekend, terwijl de ideeën en kansen die hier ontstaan van grote waarde kunnen zijn. Beide groepen zijn interessant, ondanks de verschillen, en juist vanwege de verschillen is niet het uitsluiten maar het maken van onderscheid op basis van afhankelijkheid belangrijk. Met name de onafhankelijke startup zal veel baat hebben wanneer het beter zichtbaar wordt voor publieke leiders of andere potentiële samenwerkende partijen, het kan worden betrokken bij regionale beleidsplannen en wanneer aanvullende ondersteuning en toegang tot middelen beschikbaar kunnen worden gesteld. Hieronder valt naast directe ondersteuning ook het indirect beter kunnen sturen van deze groep mensen en ideeën.

2.2.6 Nut van startup-tellingen

Zoals aangegeven in de probleemstelling en de theorie over het meten van innovatie, zegt het tel-

len van aantallen startups binnen een sector niet in hoeverre het gaat over daadwerkelijk groeipotentieel of innovatie. Idealiter zou onderzoek naar groeipotentieel daarom longitudinaal moeten zijn om zowel in- als output te meten. Zo beweren Roundy, Brockman, & Bradshaw (2017) dat veel tellingen van startups het ongelijke belang van actoren en processen in de loop van de tijd onvoldoende onderkennen. Dit neemt niet weg dat tellingen niet zinnig zijn, het kan worden gezien als eerste stap die weliswaar niet het meest nauwkeurig maar wel relatief efficiënt is. Voor interpretatie van dergelijke metingen, blijft het van belang dat criteria worden gebruikt die zo nauwkeurig mogelijk groeipotentieel meten. Vervolgens moet worden onderkend dat de getelde startups slechts een deel van de puzzel zijn en in werkelijkheid sterk kunnen verschillen in hun bijdrage aan de regionale economische dynamiek.

Ook Fazio, et al. (2016) constateren dat veel benaderingen om startups te meten los staan van de verschillen in groeipotentieel. De auteurs verwijzen naar het feit dat startups binnen een bepaalde sector regelmatig als gelijk worden beschouwd, of dat er een selectie wordt gemaakt van een resultaat, zoals de waarde van verkregen (duf) kapitaal. In het laatste geval wordt niet het aantal, maar het groeitempo van de startups als uitgangspunt genomen. Deze op kwantiteit en prestatie gebaseerde benaderingen bieden geen duidelijk zicht op eventuele vertekeningen van de data. Bij de eerste benadering is niet duidelijk of het aantal startups binnen een sector een weerspiegeling is van het daadwerkelijke groeipotentieel. De tweede benadering kijkt pas achteraf naar het moment van oprichting, waarbij de factoren die achteraf voor succes lijken te hebben gezorgd (zoals het relatieve aanbod van risicokapitaal, regionale ecosysteemeffecten of geluk) weinig zeggen over het oorspronkelijke groeipotentieel aanwezig in de regio.

Wanneer aanvullende kenmerken zoals de eerder besproken mate van afhankelijkheid bekend zijn, kan dit van voorspellende waarde zijn om het innovatie- of groeipotentieel enigszins in te schatten. Kwalitatief maatwerk blijft echter noodzakelijk voor de precieze bepaling aangezien innovatie een verzamelterm is voor een breed kwalitatief concept dat nogal eens verwarrend blijkt te zijn. Niet alle innovaties zijn hetzelfde, en daarom worden ze vaak onderverdeeld in typologieën om bepaalde kenmerken of gradaties van elkaar te onderscheiden. Het literatuuronderzoek van Garcia & Calantone

(2002) laat zien dat een overvloed aan typologieën is toegewezen aan het innovatieproces. Dit heeft ertoe geleid dat dezelfde typologie wordt gebruikt voor verschillende soorten innovaties en dezelfde innovatie wordt ingedeeld onder verschillende typologieën. De moeilijkheid lijkt voort te komen uit de verschillende actoren en processen die samen in de loop van de tijd leiden tot nieuwe innovaties. Hierbij is het niet altijd duidelijk of mogelijk om te bepalen welke actoren het meest innovatief zijn of de belangrijkste invloed uitoefenen op het proces.

Ter illustratie, veel bedrijven nemen een innovatiestrategie aan waarbij bestaande producten of technologieën worden nagebootst en verbeterd. Dit wordt door Miles, et al. (1978) omschreven als de '*analyzers*' strategie. Zelfs succesvolle *analyzers*, waaronder veel gevestigde bedrijven zoals Microsoft, worden door concurrenten vaak gezien als geweldige imitators en niet erg innovatief. Een hoogst innovatief product betekent dus niet automatisch dat het zeer innovatieve bedrijven zijn (Garcia & Calantone, 2002). Baumol (1996) gaat zelfs een stap verder door te stellen dat sommige ondernemers of startups een parasitair bestaan leiden dat schadelijk is voor de economie.¹⁷ Volgens de redenering van Baumol (2010) zijn zogenaamd innovatieve startups wat betreft hun strategie niet homogeen. Startups kunnen volgens hem worden gegroepeerd in een groot aantal dat alleen bestaande ideeën kopieert, een klein deel dat kleine innovaties introduceert, en slechts een paar Schumpeteriaanse nieuwe bedrijven die baanbrekend zijn.

Wat de interpretatie betreft, zou niet alleen het innovatie potentieel, maar ook een gebrek aan potentieel interessant kunnen zijn. Innovatiebeleid is relevant wanneer publieke leiders problemen, waarvan zij willen dat innovatie een bijdrage levert aan de oplossing, goed kunnen overbrengen. Een effectief innovatiebeleid is een beleid dat ook richting geeft aan de innovatie-inspanningen van startups. Het is bekend dat ondersteuning en subsidies vanuit publieke leiders positieve effecten kunnen hebben op R&D-investeringen van met

¹⁷ Lucas & Fuller (2017) pleiten zelfs voor een intensievere discussie over de regels binnen de institutionele context om alleen ondernemingsactiviteiten te stimuleren die daadwerkelijk meer sociale waarde creëren dan elk ander relevant alternatief. In de huidige visie kan de vraag of een ondernemingsinspanning maatschappelijke waarde creëert alleen worden gemeten aan de hand van de alternatieven die de ondernemer ter beschikking staan.

name de kleine bedrijven (Castellacci en Mee Lie, 2015). De bredere maatschappelijke effecten wat betreft de output aan innovatie, productiviteit en banen zijn echter veel minder zeker (Larédo et al., 2016; Mohnen et al., 2017). Tellingen en interpretaties van startup-activiteiten kunnen een bijdrage leveren aan verbeterde beeldvorming en mogelijk betere afstemming tussen probleem en oplossing.

Voor de bepaling van het groei- of innovatie potentieel van startups, ligt een systeem- of contextuele benadering meer voor de hand gezien het toenemende besef dat individuele startups slechts een deel van de puzzel naar succesvolle innovatie zijn. Dit sluit aan bij bibliometrisch bewijsmateriaal waaruit blijkt dat het gebruik van de term

‘entrepreneuriaal ecosysteem’ andere concepten heeft ingehaald (Malecki, 2018). Het concept van een entrepreneuriaal ecosysteem benadrukt hoe groei en innovatie mogelijk wordt gemaakt door een uitgebreide reeks aan middelen en actoren, die samen een belangrijke rol kunnen spelen bij de ontwikkeling van startups (Thornton & Flynn, 2003; Stam, 2014; Malecki, 2018). Ook het Planbureau voor de Leefomgeving (2017) benadrukt het toenemende belang van een systeembenadering in recent beleid. De (geografische) interpretatie van startup tellingen kan daarom het beste plaats vinden vanuit het perspectief van een regionaal innovatiesysteem. Dit concept wordt samen met de geografische theorieën in het volgende en laatste deel van dit theoretisch kader besproken.

2.3 Het Ondernemende Ecosysteem

Het concept ‘entrepreneuriale of ondernemende ecosysteem’ refereert aan de interactie tussen verschillende institutionele en individuele stakeholders die ondernemerschap, innovatie en groei van het midden- en kleinbedrijf bevorderen (Mazzarol, 2014; Malecki, 2018). Deze interactie moet leiden tot een gunstig klimaat waarmee ondernemerschap wordt gestimuleerd. De exacte definities van dergelijke ecosystemen worden in de literatuur op verschillende manieren geformuleerd. De gehanteerde definities zijn onder meer afhankelijk van het gekozen schaalniveau, onderzoeksdoel en de beschikbare gegevens (Malecki, 2018). In de meeste gevallen verwijst het ondernemende ecosysteem naar een combinatie of interactie van elementen die gedeelde culturele waarden produceren om ondernemersactiviteiten te ondersteunen. Dit betekent dat zowel de opkomst als de verdere ontwikkeling van innovatieve startups terug te voeren zijn op de mogelijkheden van het ecosysteem.

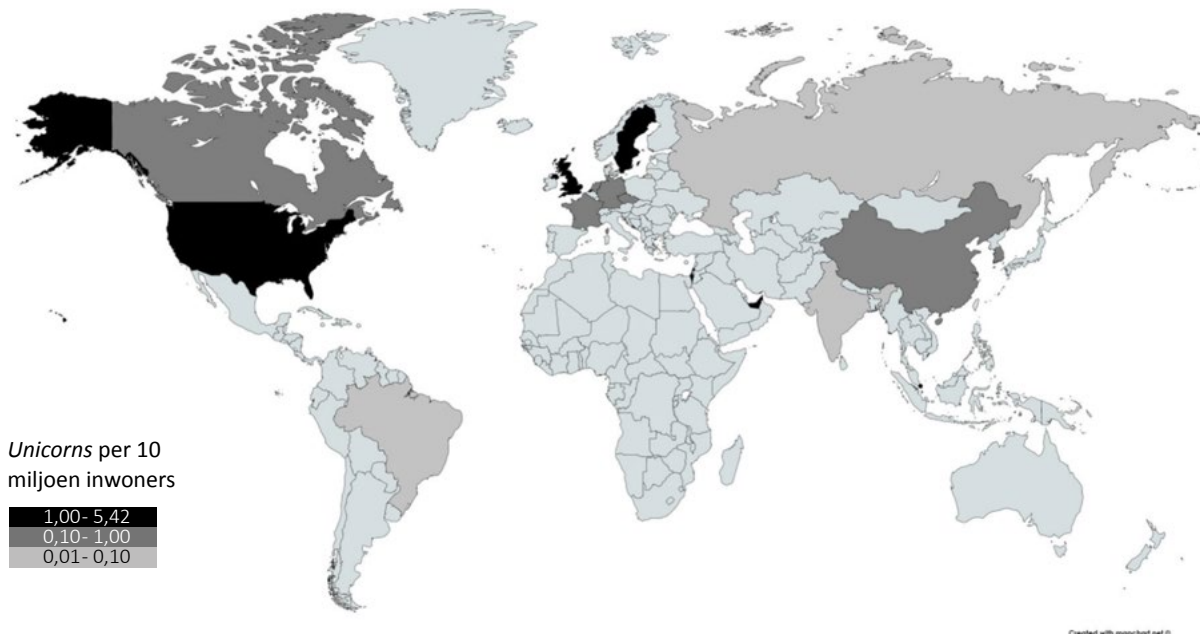
Het belang van het regionale innovatiesysteem ten behoeve van dit onderzoek ligt niet in het optimaliseren van het ecosysteem, maar in het analyseren, verklaren en betekenis geven aan de gevonden startups in de regio Noord-Nederland. Zoals gesteld in de inleiding is onvoldoende duidelijk wanneer welk concept bruikbaar is voor welke geografische regio. Dit deel, bespreekt de huidige inzichten die wel belangrijke implicaties bieden voor de analyse van de te bestuderen regio Noord-Nederland. Naast het bibliometrische overzicht van Malecki (2018) biedt een nog recentere studie van Szerb, Lafuente, Horváth, & Páger (2018) aanvullende inzichten in de relevantie van zowel de

kwantiteit als de kwaliteit van startups voor de regionale economische dynamiek. Laatstgenoemde auteurs dichten een modererende rol toe aan het ondernemende ecosysteem.

Het concept van een ondernemend ecosysteem is in theorie een allesomvattende benadering en kan daarom ook een eindeloos aantal componenten omvatten. Omdat de gegevens gebruikt door Szerb et al. (2018) niet openbaar beschikbaar zijn is het niet mogelijk om het ecosysteem van Noord-Nederland op een vergelijkbare manier te indexeren. Wel kan op basis van verschillende literatuurbronnen een inschatting worden gemaakt van het type ecosysteem, en de modererende werking. Aan de hand van de literatuur worden in de volgende paragrafen verschillende concepten en voorbeelden van regionale innovatie systemen gedestilleerd tot bruikbare implicaties voor de regio Noord Nederland. Verschillende schaalniveaus en benaderingen zullen hierbij de revue passeren. Vervolgens worden op basis van voorgaande besprekingen verschillende toepassingen op regionale entiteiten besproken, daaruit kunnen relevante bevindingen – zoals voorspellende factoren en andere mogelijke verklaringen – worden gehaald ten behoeve van innovatieve startups in het onderzoeksgebied Noord-Nederland.

2.3.1 Nationale identificatie, regionaal speelveld, wereldwijde banden

Ondanks het lokale karakter van ondernemerschap, worden ondernemende ecosystemen vaak op nationaal niveau geïdentificeerd (Brown & Mason, 2017; Malecki 2018). Het merendeel van



Figuur 2.5 Het voorkomen van Unicorns over de wereld, per 10 miljoen inwoners. Bron: (Stam & Saberi, 2017)

de Angelsaksische literatuur over ondernemende ecosystemen heeft betrekking op de VS of andere grote westerse staten of steden (Gauthier et al., 2018; Szerb et al., 2018; Malecki, 2018). Figuur 2.5 illustreert met het wereldwijde voorkomen van *Unicorns*¹⁸ de typische identificatie van innovatie op nationaal niveau, ondanks het feit dat veel van deze startups in veel kleinere geografische gebieden ontstaan: De ecosystemen rond Silicon Valley en New York bijvoorbeeld, brengen een onevenredig groot deel van *Unicorns* in de VS voort en functioneren in de praktijk als zelfstandige startup ecosysteem.

De mondiale geografie van *Unicorns* is tevens illustrerend voor een andere bevinding uit de literatuur: startup-activiteiten in minder ontwikkelde gebieden komen niet alleen minder vaak voor, maar presteren ook veel minder goed (Van Stel, Carree, & Thurik, 2005; Acs, 2006). Ecosysteemfactoren kunnen variëren, maar regio's in hetzelfde land hebben te maken met vergelijkbare nationale invloeden die de individuele beslissing voor zelfstandig ondernemerschap kunnen begunstigen of juist ontmoedigen. Voor Noord-Nederland betekent dit dat het tot op zekere hoogte vergelijkbaar is met andere westerse contexten.

Nederland is een typisch voorbeeld van een ondernemend ecosysteem op nationale schaal (Stam, 2014) en vergelijkbaar met grootstedelijke ecosys-

¹⁸ *Unicorns* zijn snelgroeiende startups met een waarde van meer dan 1 miljard dollar (nog voordat ze de beursgang bereiken).

temen in andere landen. Het presteert in vergelijking tot andere landen bovengemiddeld tot goed.¹⁹ Maar ondanks identificatie van een ecosysteem op nationaal niveau, zijn het de grootstedelijke agglomeraties die vaak in het middelpunt van de belangstelling staan. Dit is het schaalniveau waarop de meeste interacties plaatsvinden, en misschien ook wel het beste de geografische vergelijking kan worden getrokken. De straal van 100 kilometer rondom een epicentrum gehanteerd door Gauthier et al. (2018), heeft echter tot gevolg dat een groot deel van Noord-Nederland – waaronder het Noord-Nederlandse epicentrum Groningen – geen deel uitmaakt van het Nederlandse ecosysteem. Groningen is hemelsbreed meer dan 100 kilometer van Amsterdam verwijderd en valt daarmee ver buiten de geografisch gedefinieerde driehoek Amsterdam-Den Haag-Eindhoven zoals gehanteerd in het onderzoek van Gauthier et al. (2018).

Een rapport van de Organisatie voor Economische Samenwerking en Ontwikkeling (OESO) (Marsan & Maguire, 2011) gebruikt een regionale indeling vergelijkbaar met dit onderzoek en splitst Nederland in vier delen (Noord, Oost, Zuid, West). Het onderzoek classificeert regio's op basis van traditionale innovatie indicatoren en aanvullende

¹⁹ Nederland presteert goed als het gaat om regionale innovatiegerelateerde kenmerken. Dit blijkt uit verschillende bronnen waaronder de indexrapporten in tabel 2.2, een verklaring van de Rijksoverheid (2018) en uit een door Marsan & Maguire (2011) samengesteld rapport voor de Organisatie voor Economische Samenwerking en Ontwikkeling (OESO) over regionale innovatiegerelateerde (en sociodemografische en economische) kenmerken.

sociodemografische en economische kenmerken. Zuid-Nederland behoort tot de zogenaamde *Knowledge and technology-hubs*. Dit zijn de regio's met de beste prestaties op gebied van wetenschap en technologie (vooral gebaseerd op R&D-indicatoren en patenten) en met een industriële structuur waarvan een aanzienlijk deel van de productie uit hoogtechnologische sectoren voortkomt. De overige Nederlandse regio's, waaronder Noord-Nederland, vallen onder de zogenaamde *Service and natural resource regions in knowledge-intensive countries*. Deze categorie omvat regio's die in eigen land op gebied van innovatie vaak als tweederangs worden gezien. Ze zijn over het algemeen minder dichtbevolkt, maar ondanks de tweederangs status is er sprake van een goed ontwikkelde opnamecapaciteit van kennis vanwege een hoog opgeleide beroepsbevolking.

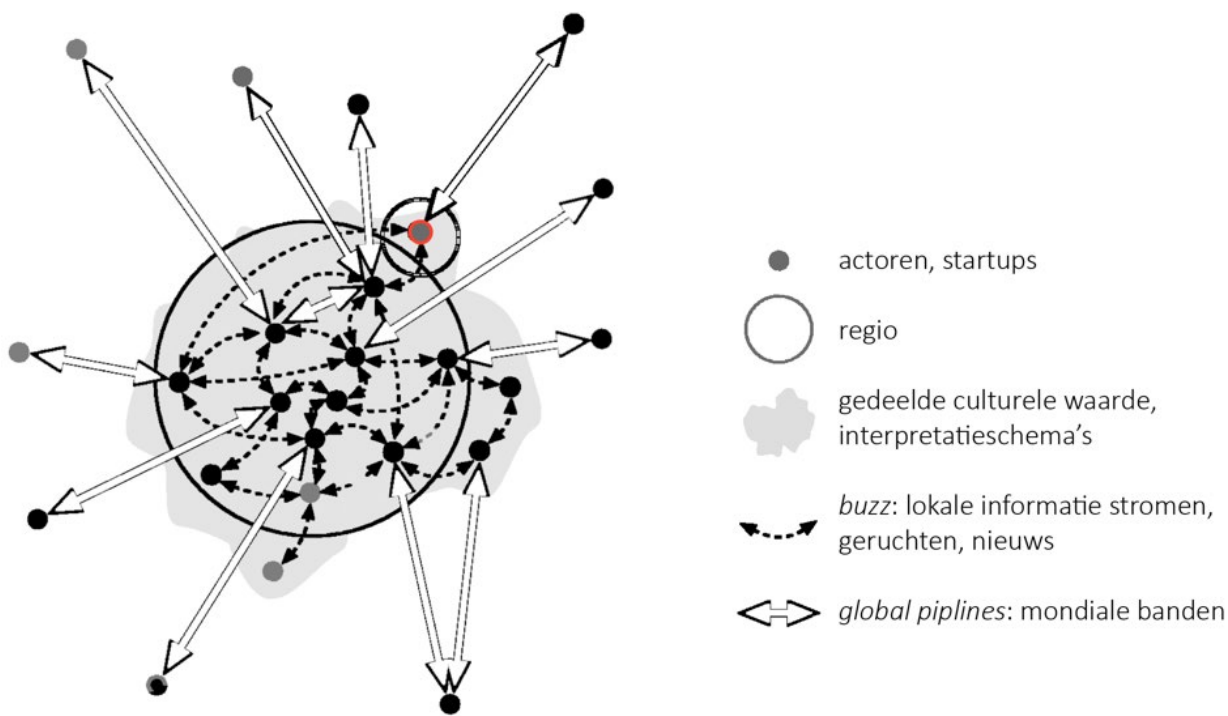
Andere onderzoeken hanteren een kleinere straal, vaak gebaseerd op een enkele stad of district. Zo bestudeerd het door Malecki (2018) aangehaalde onderzoek van Cukier, Kon, & Lyons (2016) het ecosysteem van New York. Hierbij is een straal van 50 kilometer gehanteerd. Dit is zowel wat betreft territoriale schaal als westerse context vergelijkbaar met de regio Noord-Nederland. Natuurlijk mag er geen directe vergelijking worden getrokken omdat de Noord-Nederlandse regio met veel minder status en bekendheid een inferieure rol speelt in vergelijking tot dat van de wereldwijd sterker verbonden metropool New York. Ook volgens de OESO functioneert Noord-Nederland niet als wereldwijde kennishub (Marsan & Maguire, 2011). Wereldwijde verbondenheid is volgens Gauthier et al. (2018) echter wel cruciaal om startups kansrijk te laten zijn. Ook in de academische literatuur is men het er over eens dat elk ondernemend ecosysteem, ongeacht het schaalniveau en dus inclusief een regio als Noord-Nederland, tot op zekere hoogte wereldwijde banden moet hebben (Malecki, 2018).

Voor het blootleggen van deze mondiale banden zijn verschillende benaderingen gebruikelijk. Een van de manieren is om het wereldwijde marktgebied van een ecosysteem te meten. Gauthier et al. (2018) kijken hiervoor naar het gemiddelde percentage klanten buiten het land of continent. Startups in Europa kunnen bijvoorbeeld verkopen aan buitenlandse klanten in andere Europese landen, maar om geografische voorkeuren te voorkomen wordt ook gekeken in hoeverre ze buiten Europa verkopen.

Een andere benadering betreft het kwantificeren van betekenisvolle relaties die ontstaan tussen oprichters van startups afkomstig uit verschillende ecosystemen. Hierbij wordt gekeken naar het aantal kennisrelaties tussen oprichters dat ontstaat als gevolg van het reizen naar een ander ecosysteem. Gauthier et al. (2018) richten zich met name op de verbindingen met 's werelds topecosystemen. De achterliggende gedachte is dat deze ecosystemen een voortrekkersrol spelen in de wereldwijde structuur van kennis, ideeën, mensen en organisaties.

Naast wereldwijd marktgebied en interactie tussen oprichters van startups uit verschillende (top) ecosystemen, kan de mate waarin innovaties naar verwachting wereldwijd toonaangevend zullen zijn, wijzen op de aanwezigheid van wereldwijde banden. Gauthier et al. (2018) gebruiken verzamelde statistieken over het percentage startups in een ecosysteem dat zegt een geheel nieuw product voor wereldwijde markten te ontwikkelen. Een dergelijke statistiek zegt iets over de ambities van oprichters, en wordt in verband gebracht met de aanwezigheid van mondiale banden in een ecosysteem. In welke mate startups profiteren van wereldwijde kennis, is namelijk niet alleen bepalend voor het kiezen van een nieuw businessmodel of het ontwikkelen van een nieuwe technologie, maar ook bepalend om er zeker van te zijn dat er daadwerkelijk sprake is van iets nieuw op de markt. De aanwezige kennis zal ertoe leiden dat mensen met meer overtuiging en ambitie, eerder geneigd zullen zijn om startups op te zetten, met als gevolg dat deze ook meer kans hebben succesvol te zijn.

Wat betreft het meten van mondiale banden volstaat het bespreken van voorgaande benaderingen zoals toegepast door Gauthier et al. (2018). Binnen de scope van dit onderzoek is het onlogisch dieper op deze kwestie in te gaan aangezien het slechts de bedoeling is een algemeen beeld te verkrijgen. Bovendien wordt Noord-Nederland niet tot de zogenaamde topecosystemen met hoogstaande mondiale banden gerekend. Wat overigens niets af doet aan het feit dat wereldwijde banden aanwezig laat staan belangrijk zullen zijn in en voor Noord-Nederland. Dit weerspiegelt ook de constatering van Bathelt, Malmberg, & Maskell (2004) dat mondiale banden – afgezien van de lijnen tussen bedrijven en afzetmarkt – een belangrijke rol spelen bij de ontwikkeling van startups in een ondernemend ecosysteem.



Figuur 2.6 De structuur en dynamiek van lokale buzz en wereldwijde pijpleidingen. Overgenomen en bewerkt uit: Bathelt et al. (2004, p.46).

Naast mondiale banden – of *global pipelines* om in de bewoording van Bathelt et al. (2004) te spreken – is er ook expliciete erkenning en consensus over het belang van lokale banden. Zo wordt door Gauthier et al. (2018) lokale verbondenheid in een ecosysteem sterk geassocieerd met hogere startup prestaties. Deze lokale verbondenheid wordt door hen wederom gemeten door te kijken naar het aantal betekenisvolle relaties met andere oprichters, startups en actoren in het ecosysteem, en het gemak waarmee deze relaties kunnen worden aangegaan. Het cruciale belang van lokale wisselwerking tussen startups onderling (en andere actoren binnen een regio of ecosysteem) wordt ook in de academische literatuur bevestigd (Malecki, 2018). Noemenswaardig is de verwijzing van Bathelt et al. (2004) naar het bestaan van hoogwaardige en relevante lokale informatiestromen aangeduid als ‘buzz’:

Buzz refers to the information and communication ecology created by face-to-face contacts, co-presence and co-location of people and firms within the same industry and place or region. This buzz consists of specific information and continuous updates of this information, intended and unanticipated learning processes in organized and accidental meetings, the application of the same interpretative schemes and mutual understanding of new knowledge and technologies, as well as shared cultural traditions and habits within a particular technology field, which stimulate the establishment of conventions and

other institutional arrangements. (p.38

Deze lokale ‘buzz’ laat ondernemersecosystemen uitgroeien tot een succesvol klimaat voor startups, vaak gekenmerkt door een typische lokale cultuur (Malecki, 2018). Naast oprichters en andere actoren in de private sector dragen ook non-profitorganisaties en sociale ondernemers (Birch & Whittam, 2008) bij aan een gunstig klimaat.

Daarnaast worden de kennisinstellingen, naast de ondernemers zelf, regelmatig genoemd als zeer prominente actor in ondernemersecosystemen en een groot deel van de literatuur richt zich dan ook op universiteiten als knooppunten van zowel lokale als mondiale banden. De aanwezigheid van kennisinstellingen leidt mogelijk tot meer opstartactiviteiten. Voor Noord-Nederland bevestigd dit dat Groningen (met een Universiteit en Hogeschool) en in mindere mate Leeuwarden (Hogeschool) als epicentra van de regio kunnen worden gezien.

Deze paragraaf sluit af met de conclusie dat toegang tot zowel lokale als meer gespecialiseerde kennis en middelen die niet lokaal beschikbaar zijn, als bijzonder belangrijk worden beschouwd in de vroege stadia van startupvorming. De verschillende schaalniveaus die in deze paragraaf zijn besproken komen samen in figuur 2.6. Volgens de oorspronkelijke makers moet de visualisatie een (kennis) cluster voorstellen dat regionale grenzen overstijgt. Het kan echter ook toepasbaar zijn op de werking van een ondernemend ecosysteem waarin mondi-

ale en lokale banden elkaar versterken. Het grijze gebied zou bijvoorbeeld Nederland – en tevens het ecosysteem zoals gedefinieerd door Stam (2014) – kunnen voorstellen, terwijl Gauthier et al. (2018) de grote cirkel als afbakening zouden hanteren voor het Nederlandse ecosysteem. De kleinere cirkel zou dan Noord-Nederland zijn, waarbinnen startups en actoren onderling verbonden kunnen zijn, maar ook profiteren van mondiale banden die ze zelf onderhouden of die hen indirect via lokale informatiestromen weet te bereiken. Startups en actoren gevormd in en rondom de kennisinstellingen van Groningen zouden het knooppunt van kennis in de regio kunnen vormen. In hoeverre en onder welke voorwaarden geografische nabijheid van kennis van belang is wordt nader besproken in de afsluitende volgende paragrafen.

2.2.2 Kennisontwikkeling en het sub-urbane landschap

Glaeser (2011) is in het inleidende woord al aan bod gekomen als fervent voorstander van kennis clusters in hoge dichtheden. Daarmee is hij zeker niet de enige voor wie het ruimtelijke agglomeratieproces een bron van grote fascinatie is geweest. In het theoretische overzicht van Mason & Brown (2017) worden door de tijd verschillende perspectieven geïdentificeerd die ruimtelijke agglomeratieprocessen door de tijd associëren met innovatie. De perspectieven komen soms sterk overeen met het gedachtegoed van de evolutionaire (economische) geografie. Na een korte introductie van deze benadering, worden twee belangrijke ruimtelijke agglomeratietheorieën – waarvan Marshall (1890) en Jacobs (1961) worden beschouwd als grondleggers – nader besproken.

Evolutionaire geografie is volgens van Rietbergen, Brinkman, Kuypers, Rammos, & Witte (2008) het beste te omschrijven als een theorie van economische veranderingen en aanpassingen waarbij de volgende aspecten worden geanalyseerd:

- hoe nieuwe variatie in de economie terechtkomt en hoe deze leidt tot verandering van technologie, marktstructuur, sectorstructuur en instituties,
- hoe de selectieomgeving, die bestaat uit zowel markten als instituties, bepaalt welke variatie overleeft en hoe dit leidt tot een bepaalde mate van stabiliteit (*lock-in*²⁰),

²⁰ Negatieve *lock-in* wijst op een gebrek aan aanpassingsvermogen in een regio.

- hoe radicale innovaties in technologie en organisatievormen samengaan met veranderingen in de selectieomgeving zelf, in een proces van co-evolutie,
- hoe veranderingsprocessen zich onomkeerbaar en padafhankelijk kunnen voltrekken, en hoe economisch gedrag suboptimale²¹, onvoorspelbare en meervoudige uitkomsten kan opleveren. (p.29)

Evolutionaire economische geografie en innovatie zijn onlosmakelijk verbonden. Daarbij lijkt de evolutionaire benadering de nadruk te leggen op de structurele en lange termijn veranderingen. Tegelijkertijd verschaft de benadering ook inzicht op korte termijn veranderingen: lange termijn veranderingen voltrekken zich stapje voor stapje en radicale veranderingen vinden vrij plotseling plaats. De theorie heeft zich ontwikkeld tot een samenhangend analysekader voor het bestuderen van innovatie (Frenken, 2011). Ideeën uit de evolutionaire benadering keren impliciet of expliciet terug tijdens de verdere besprekingen in dit hoofdstuk, maar worden uiteindelijk (in paragraaf 3.3) ondervangen door het ondernemende ecosysteem als hoofdbenadering van dit hoofdstuk.

Zoals aangekondigd worden nu twee belangrijke perspectieven van het ruimtelijke agglomeratieproces geïntroduceerd. Het gaat om het Marshalliaanse agglomeratie-effect van specialisatie (Marshall, 1890) en de tegenstrijdige gedachte over diversificatie van externaliteiten (Jacobs, 1961). Beide perspectieven worden in eerste instantie grotendeels genegeerd of niet opgemerkt, en trekken pas later meer aandacht (Nowlan, 1997). Ook Glaeser raakt geïnspireerd en geeft beide perspectieven een prominente plaats in het door hem geleide onderzoek *Growth of Cities* (Glaeser, Kallal, Scheinkman, & Shleifer, 1992).

Het eerstgenoemde Marshalliaanse perspectief is door Glaeser et al. (1992) verenigd met aanverwante opvattingen, en vervolgens het *MAR-spillover* genoemd. Volgens dit *spillover*-perspectief verbeterd regionale specialisatie naar een bepaalde industrie de innovativiteit binnen die sector. De theorie baseert zich op het idee dat de nabijheid van bedrijven binnen vergelijkbare industrieën de uitwisseling van kennis en ideeën het beste stimuleert. Hoe

²¹ Zie ook paragraaf 2.2.6 waar Baumol (1996) wijst op het parasitaire bestaan van sommige ondernemingen met als aanvulling voetnoot 17.

dichter bedrijven in een gespecialiseerd cluster bij elkaar liggen, hoe groter de *MAR-spillover* en hoe groter de kans op innovatie. Dit Marshalliaanse perspectief is nauw verwant aan andere regionale specialisatie theorieën. Alhoewel veel van deze theorieën slechts productiviteitsverhoging nastreven, benadrukken andere clustertheorieën ook de positieve impact op innovatieve activiteiten als gevolg van diepgaande en gespecialiseerde kennis. Hieronder valt ook de clustertheorie van Porter (1998a; 1998b), die naast hogere productiviteit en innovativiteit ook beweert dat sectorale clustering en competitie de opkomst van nieuwe startups in de regio stimuleert.

Volgens het alternatieve *spillover*-perspectief van Jacobs, is de nabijheid van bedrijven uit verschillende complementaire in plaats van vergelijkbare industrieën juist doorslaggevend. Jacobs (1961) beargumenteert dat interactie van bedrijven en economische actoren met uiteenlopende activiteiten, het zoeken en experimenteren ten behoeve van innovatie vergemakkelijken. De aanwezigheid van verschillende industrieën maakt het mogelijk dat ideeën die door één industrie zijn ontwikkeld, in andere sectoren kunnen worden opgepikt en doorontwikkeld. Een gediversifieerde regionale ondernemerslandschap zal, volgens dit perspectief, de kennis van het individuele bedrijf vergroten. In regionaal opzicht wordt dit fenomeen geduid als diversificatie van externaliteiten met als positief effect het bevorderen van de regionale innovativiteit. De zogenaamde 'Jacobs-externaliteiten' kunnen worden omschreven als schaalvoordelen buiten het bedrijf die leiden tot innovatieve werkwijzen en gedrag, zoals de ontwikkeling van cross-sectorale verbanden en nieuwe combinaties van bestaande technologieën en materialen. Deze nieuwe combinaties zouden ook het ontstaan van nieuwe startups stimuleren.

Ondanks vele inspanningen om bovengenoemde theorieën tegen elkaar af te zetten, biedt de literatuur geen uitsluitsel over welke van de twee perspectieven het meeste voor regionale innovatie of startup-formatie pleit (Van der Panne & Van Beers, 2006). Toch benoemen Desrochers & Hospers (2007) en De Groot, Poot, & Smit (2007) dat een meerderheid van de artikelen aan de kant van het laatst genoemde perspectief van Jacobs is komen te staan. Een enigszins nauwkeurige vergelijking is echter problematisch, dit heeft te maken met het beperkte regionale beeld van innovatie als gevolg van ontoereikende methoden (zoals besproken in

deelhoofdstuk 2.1), maar ook met de moeizame ontwikkeling van een werkbaar concept van diversiteit en specialisatie – daarover straks meer. De onbesliste strijd tussen de twee perspectieven is niet het probleem. Regionale specialisatie hoeft niet ten koste te gaan van industriële diversiteit of andersom. Studies tot nu toe ondersteunen niet alleen beide krachten, maar erkennen ook dat ze (goed) naast elkaar (kunnen) bestaan. Desrochers & Sautet (2008) concluderen hierover:

In any case, if one recognizes some useful role for geographical proximity in terms of generating innovative behavior, then a good regional context for innovation would seem to be a diversified city made up of many specialized clusters—which is historically what most thriving cities have spontaneously developed into. (p.827)

Deze opvatting sluit aan bij het ideaal dat Glaeser (2011) voor ogen heeft. Maar daarmee is het laatste woord nog niet gezegd. Want ondanks het feit dat de twee perspectieven geschikt zijn voor analytische doeleinden, suggereren zowel de ontwikkelingen binnen de evolutionaire geografie als het atypische onderzoeksgebied dat uitgangspunt is van deze studie, dat een geschikter genuanceerder perspectief nodig is. Dit perspectief zou ook inzicht moeten bieden in meer rurale omgevingen in plaats van de vele agglomeratietheorieën gebaseerd op kennisinsentieve (wereld-)steden en toecosystemen. De volgende paragrafen bespreken eerst de ontwikkelingen vanuit de evolutionaire geografie en vervolgens het rurale perspectief van Noord-Nederland.

De problematische operationalisering van regionale diversiteit en specialisatie komt nu aan de orde. Hierbij is de belangrijkste ontwikkeling dat

Kader 2.3 Verwarring rondom specialisatie

Nog los van de discussie of de aard van een kennisuitwisseling toebehoort aan gerelateerde diversiteit of specialisatie, zijn de domeinen waarin een land of regio zich specialiseert vaak zo talrijk en breed geformuleerd dat het de suggestie van specialisatie overtreft (Jacobs, 2000; Janssen, 2015). Bovendien is er sprake van veel overlap tussen de verschillende domeinen waarin regio's zich (wensen te) specialiseren (Asheim, Boschma, & Cooke, 2011).

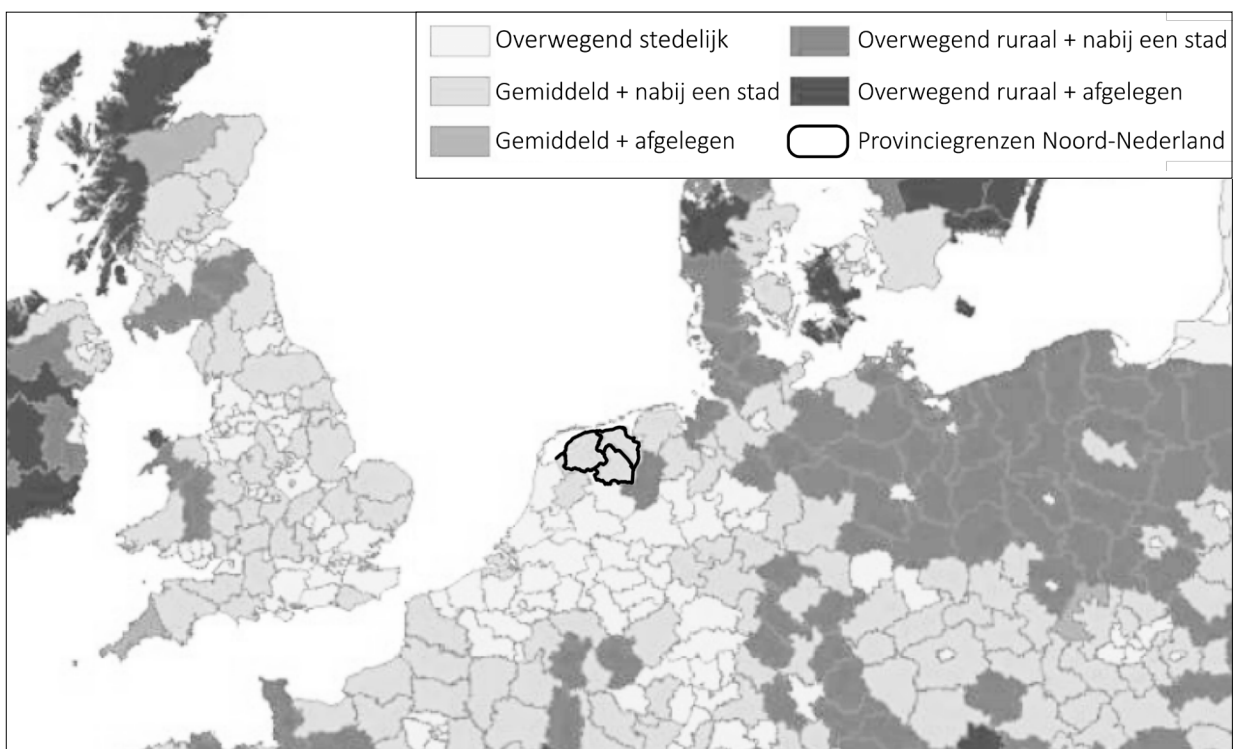
ruimtelijke analyses van industriële clusters, inclusief de studie van Glaeser et al. (1992), met name gericht zijn geweest op de ruimtelijke concentratie van afzonderlijke sectoren; meer recentelijk, zijn ze zich ook gaan richten op gerelateerde industrieën binnen de sectoren zelf (Frenken, Van Oort, & Verburg, 2007; Janssen, 2015). De achterliggende gedachte hierbij is dat, afhankelijk van de sectorale indeling, sprake kan zijn van veel diversiteit binnen de afzonderlijke sectoren (zie ook paragraaf 2.1.8). Deze interindustriële kennis staat bekend als *related diversity* of gerelateerde diversiteit.

Sommige studies suggereren of constateren een positief effect van gerelateerde diversiteit op innovatie of startup-formatie in westerse landen (Mowery, Oxley, & Silverman, 1996; Boschma & Lammarino, 2009; Boschma, Balland, & Kogler, 2014). Sterker nog, gerelateerde diversiteit zou tot meer innovatie leiden dan niet-gerelateerde diversiteit. Echter geldt voor disruptieve innovatie mogelijk dat juist ongerelateerde diversiteit van groter belang is (Nootboom, Van Haverbeke, Duysters, Gilsing, & Van den Oord, 2007; Enkel & Gassmann, 2010; Castaldi, Frenken, & Los, 2014). Hoe dan ook, implicaties van studies over regionale diversiteit en specialisatie zijn alleen bruikbaar wanneer de sectorale indeling, en daarbinnen de gehanteerde definitie van diversiteit of specialisatie, in overweging worden genomen. De opkomst van dezelfde innovatie of startup kan zowel worden toegeschreven aan gerelateerde diversiteit als het

concept van specialisatie. Dit is volledig afhankelijk van de operationalisering, en werpt een nieuw licht op de Jacobs- en MAR-spillover theorieën (zie ook kader 3.1. voor aanvullende verwarring rondom specialisatie in sectoren).

Wat betreft de toenemende aandacht voor (on)gerelateerde diversiteit, stelt Janssen (2015) dat de kans dat actoren binnen een economie innovatieve oplossingen vinden om een sterke positie van een industrie te behouden toeneemt wanneer het kennis vanuit ongelijksoortige industrieën implementeert. Echter wordt ook onderkend dat kennisuitwisseling moeilijk is wanneer partijen cognitief uit elkaar liggen (Nootboom, 2000). Dit wordt ook bevestigd door Boschma & Lammarino (2009). Zij concluderen in hun studie naar Italiaanse provincies, dat de cognitieve nabijheid niet te klein of te groot moet zijn om leermogelijkheden bij te laten dragen aan economische groei. Frenken et al. (2007) komen in hun studie binnen de Nederlandse economie tot een soortgelijke conclusie, terwijl Boschma et al. (2014) op basis van Noord-Amerikaanse patentgegevens aantonen dat een hoge mate van gerelateerde diversiteit geassocieerd kan worden met een hogere mate van innovatie.

De bovenstaande theorieën suggereren een strategie van slimme groei. Het is dan ook geen verrassing dat het concept van slimme groei de kern vormt van de Europese groeistrategie en onder meer tot doel heeft de bestaande innovatie-kloof tussen



Figuur 2.7 OESO regionale ruraal-stedelijkheid-typologie in Europa, met daarin de drie provincies in het onderzoeksgebied Noord-Nederland. Overgenomen en bewerkt uit: Dijkstra & Ruiz (2010, p.7).

de Europese Unie en enkele van haar belangrijkste handelspartners te verkleinen (Naldi, Nilsson, Westlund, & Wixe, 2015). De meeste ruimtelijke theorieën suggereren dat vooral grootstedelijke regio's een positieve werking hebben op startups, aangezien zij toegang verschaffen tot lokale, regionale en mondiale kennisbronnen en afzetmarkten. Het is echter minder duidelijk in welke mate het concept van slimme groei van toepassing is op meer rurale gebieden zoals in Noord-Nederland.

Er bestaan verschillende regionale typologieën die gebaseerd zijn op verschillende institutionele en economische indicatoren. Wetenschappers en beleidsmakers tonen in toenemende mate interesse in suburbane landschappen; zogenaamde tussenliggende vormen in het ruraal-stedelijke continuüm (Renski, 2014). De OESO-typologie van Dijkstra & Ruiz (2010) in figuur 2.7 kijkt niet alleen naar het percentage inwoners dat in ruraal of stedelijk gebied woont, maar ook naar de nabijheid van een stad met minimaal 50.000 inwoners. Noord-Nederland zou vanuit een Nederlandse bril als ruraal kunnen worden aangemerkt (van Dongen, Jonkeren, & Raspe, 2014), maar volgens de OESO-typologie is eerder sprake van een gemiddeld tussengebied. Copus et al. (2011) voegen indicatoren over de economische structuur en prestaties toe aan de OESO-typologie van Dijkstra & Ruiz (2010).

Hieruit blijkt dat Noord-Nederland zeker geen typische ruraal geval is, maar een tussengebied met een bovengemiddelde regionale economische prestatie en sterke secundaire en tertiaire sector. Uit een rapport van het PBL (van Dongen et al., 2014) wordt het voorgaande bevestigd in een uitgebreid onderzoek naar de determinanten van de nationale vestigingspatronen van bedrijven uit de topsectoren. Hieruit blijkt dat Noord-Nederland economisch gezien duidelijk deel uitmaakt van de Nederlandse periferie. Ondanks dit gegeven, constateren studies dat economische verschillen tussen het stedelijke en rurale gebied op sommige plaatsen vervagen (Renski, 2014; McCann & Ortega-Argilés, 2015). Tegelijkertijd, suggereren deze studies dat ondernemerschap en groeipatronen in de zogenaamde tussenliggende gebieden waar Noord-Nederland toe behoort, heel verschillend kunnen zijn ten opzichten van de stedelijke kerngebieden zoals de Randstad in Nederland.

In een onderzoek naar startups, overleving en groei in de VS laat Renski (2014) zien dat veel rurale gebieden nabij steden relatief hoog scoren bij metingen van ondernemersprestaties. Specifiek toont de studie aan dat rurale delen van metropolitane gebieden, in vergelijking tot stedelijke kerngebieden en afgelegen plattelandsgebieden, de grootste concentratie van startups hebben, evenals lagere

Tabel 2.3 Slimme (inclusief rurale of perifere) groei.

Determinant	Meeteenheid
Urbanisatie-economieën	Bevolkingsdichtheid (Jacobs, 1969)
Industriële specialisatie	Locatiequotiënt (Marshall, 1890)
Diversiteit / ongerelateerde diversiteit	Entropy (of Shannon-index) op bijvoorbeeld een tweecijferige SBI-code. (Frenken et al., 2007)
Gerelateerde diversiteit	Entropy (of Shannon-index) op bijvoorbeeld vijfcijferig SBI-code (Frenken et al., 2007)
Voorzieningen	<ul style="list-style-type: none"> - Toegang tot open ruimte en gevarieerde landschappen (Cho, Poudyal, & Roberts, 2008) - Toegang tot natuurlijke en outdoorvoorzieningen (meren, kustgebieden, natuurgebieden, enz.) (McGranahan, Wojan, & Lambert, 2011) - Gebouwde voorzieningen (musea, historische bezienswaardigheden, staatsparken, pretparken, golfclubs, skigebieden) (Deller, Tsai, Marcouiller, & English, 2001; Markeson & Deller, 2012) - Klimaatomstandigheden (Gunderson & Ng, 2006) clubs, skigebieden (Deller et al., 2001; Markeson & Deller, 2012) - Klimaatomstandigheden (Gunderson & Ng, 2006) - Climate conditions (Gunderson & Ng, 2006)
Creatieve gemeenschappen	<ul style="list-style-type: none"> - Toegang tot personeel uit de creatieve klasse (McGranahan & Wojan, 2007; McGranahan et al., 2011) - Lokale markten (bijvoorbeeld lokale festivals) - Aanwezigheid van sectoroverschrijdende activiteiten zoals toerisme, eten, drinken en culturele productie (Anwar McHenry, 2011; Bell & Jayne, 2010).
Externe kennisbanden	<ul style="list-style-type: none"> - Toegang tot hoogopgeleide personeel - Toegang tot bijvoorbeeld KIBS-medewerkers, cognitieve vaardigheden (Johansson, Johansson, & Wallin, 2015) - Toegang tot ICT, bijvoorbeeld internet van hoge snelheid - Banden tussen leidinggevende van bedrijven - Verschillende soorten samenwerking tussen bedrijven onderling (Torre & Rallet, 2005) - Wereldwijd actieve bedrijven (McCann & Acs, 2011)

Bron: Naldi, Nilsson, Westlund & Wixe (2015)

uitvalpercentages van startups in de hightech industrie. McCann & Ortega-Argilés (2015) komen in hun studie binnen de Europese context tot de conclusie dat slimme specialisatie gecombineerd met plaatsgerichte benaderingen, naar verwachting een succesvolle strategie is voor de zogenaamd tussenliggende regio's. Naldi, Nilsson, Westlund, & Wixe (2015) beargumenteren dat indicatoren gebaseerd op patenten en R&D-cijfers geografisch geclusterd en verdeeld zijn over de meest verstedelijkte regio's, wat impliceert dat traditionele en op technologie gebaseerde metingen niet of nauwelijks de innovatieve activiteiten zullen identificeren in perifere of rurale gebieden. Perifere gebieden zoals (delen van) Noord-Nederland zijn mogelijk minder aantrekkelijk voor niet-lokale bedrijven, maar goede economische prestaties inclusief een aanwas van lokale nieuwe startups zou kunnen wijzen op slimme rurale of perifere groei (zie tabel 3.1) in dit soort gebieden.

De meeste theorieën met betrekking tot de determinanten in tabel 3.1 zijn al besproken. Dat geldt deels voor urbanisatie-economieën, maar zeker voor industriële specialisatie en (on)gerelateerde diversiteit. Het belang van (externe) kennisbanden is in de vorige paragraaf besproken en in de volgende en laatste paragraaf komen de voorzieningen en creatieve gemeenschappen aan bod.

Deze paragraaf heeft ook laten zien dat theorieën over gespecialiseerde en gediversifieerde kennisclusters hebben geleerd van ideeën uit de evolutionaire benadering. Zo lijken diversiteit en specialisatie – gezien vanuit evolutionaire theorieën over gerelateerde diversiteit – soms meer op elkaar dan aanvankelijk gedacht. Uiteindelijk blijkt vooral gerelateerde diversiteit doorslaggevend voor het verklaren van innovatie en startup-formatie, terwijl ongerelateerde diversiteit moeilijker samengaat en daarom in mindere mate tot startup-formatie zal leiden. Aan de andere kant, zou juist het verbinden van ongerelateerde diversiteit interessant kunnen zijn voor radicale innovaties. Een andere bevinding is dat veel studies omtrent specialisatie en gerelateerde diversiteit vooral gebaseerd zijn op het technologische aspect en patenten in verstedelijkte regio's. Naar verwachting zal de invloed van specialisatie en diversiteit op startup-formatie per sector verschillend zijn, en wellicht ook anders zijn in meer rurale of perifere gebieden. Er is minder onderzoek dat inzicht verschaft in sub-urbane omgevingen zoals Noord-Nederland, maar de beschikbare studies suggereren dat er mogelijk al-

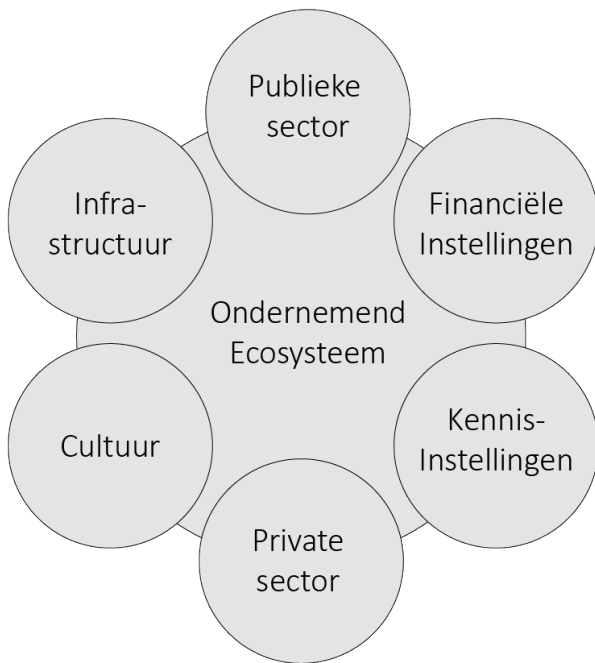
ternatieve verklaringen te vinden zijn voor de economische groei in de vorm van niet-technologische innovaties en startups.

2.2.3 Verklaringen voor startup-formatie

In de vorige paragrafen is de werking van het regionale kennis- en innovatiesysteem besproken, waarbij een aantal verklarende factoren voor innovatie- en startup-vorming al aan bod zijn gekomen. Aansluitend hierop, worden eerdere verklaringen aangevuld en samengebracht tot een overzicht van meest relevante verklaringen ten behoeve van een uiteindelijke regionale analyse op het micro-schaalniveau in Noord-Nederland. Om te beginnen, zal kort worden ingegaan op algemene locatietheorieën. Deze worden vervolgens teruggekoppeld aan de aanknopingspunten van de voorgaande paragrafen.

Locatietheorieën verklaren waarom een bedrijf voor de ene locatie kiest en niet voor de andere (Atzema, Lambooy, van Rietbergen, & Wever, 2002). Wat betreft de klassieke locatietheorieën, kan er worden aangenomen dat de optimalisatie van het productieproces, waarbij winst wordt gemaximaliseerd of kosten worden geminimaliseerd, lang niet altijd van toepassing zullen zijn op de vroege stadia van startupvorming. De welbekende zolderkamertjes, schuurtjes en tegenwoordig ook de bedrijfsverzamelgebouwen als broedplaats van startups, vormen duidelijk een uitzondering op veel klassieke locatietheorieën die stellen dat de beschikbaarheid van specifieke productiefactoren de locatiekeuze bepalen. Productiefactoren spelen wel degelijk een belangrijke rol, maar veel startups – waarvan de levensvatbaarheid nog niet is aangetoond en de zoektocht naar het beste bedrijfsmodel nog in beweging is – zullen minder gebonden zijn aan specifieke locatietekenen, zoals het geval bij de klassieke voorbeelden uit de meer traditionele industriële en agrarische sector.

Neoklassieke locatietheorieën – zoals bijvoorbeeld de centrale plaatsentheorie van Christaller (1966) – hebben de dienstensector als uitgangspunt en passen daarom beter bij een moderne economie. Maar ondanks dit gegeven, schiet de neoklassieke stroming naar verwachting tekort omdat innovatieve startups per definitie weinig te maken zullen hebben met bestaande concurrentie in de verzorging of bediening van een regio. Daarentegen bieden behaviourale theorieën – waarin de ondernemer centraal staat – wél aanvullende verklaringen voor vestigingslocaties van innovatieve startups;



Figuur 2.8 Versimpelde grafische visualisatie van een ondernemend ecosysteem. Overgenomen en bewerkt van Isenberg (2011).

nomen, komen de verschillende locatietheorieën en de ideeën uit de evolutionaire economische geografie samen in de benadering van het ondernemende ecosysteem, dat inmiddels uitgegroeid is tot een volwaardige benadering om de belangrijkste actoren, interacties en ‘cognitieve mind-sets’ (Mason & Brown, 2017; Zie ook figuur 2.9) die een verklarende rol spelen voor de het ontstaan van startups, af te bakenen²². Ondanks de conceptuele potentie van het ondernemend ecosysteem om geografische verklaringen voor startup-formatie te identificeren, is het in het kader van dit exploratieve onderzoek niet nodig of mogelijk om elk aspect en de volledige complexiteit van deze benadering te betrekken. Om het concept te versimpelen, kan het ondernemend ecosysteem worden opgesplitst in een zestal van (f)actoren (zie figuur 2.8) waarbinnen ook de locatietheorieën tot hun recht kunnen komen.

juist ook in de vroegere stadia van startup-vorming.

Bij behaviourale theorieën is de vestigingsplaats niet de uitkomst van een berekening op basis van kosten of concurrenten in de regio, maar een gevolg van persoonlijke ervaringen of omstandigheden. Zowel de sociale als economische omgeving van de ondernemer zijn dan bepalend (Louw, Needham, Holden, & Pen, 2004). Onderzoek bevestigt deze redenering. Zo vormt inbedding in sociale netwerken volgens Stam (2007) een aanzienlijke beperking voor de locatiekeuze in de vroege levensloop van een innovatieve startups. Arauzo & Manjon (2004) bevestigen dat kleine bedrijven grotendeels lijken te worden gestuurd door de voorkeuren van de persoonlijke ondernemer. Behaviourale theorieën gaan er van uit dat ondernemers eerder geneigd zullen zijn om dicht bij hun (cognitieve of fysieke) oorspronkelijke locatie (bijvoorbeeld woonadres) te blijven. Pas bij het opschalen van de onderneming zullen andere locatietheorieën een grotere rol gaan spelen.

In tegenstelling tot de eerder geïntroduceerde locatietheorieën, benadrukken institutionele locatietheorieën de invloed van maatschappelijke en culturele instituties. Institutionele locatietheorieën komen terug in de benadering van de evolutionaire economische geografie (geïntroduceerd en besproken in paragraaf 3.1.2). In feite probeert de evolutionaire economische geografie de kloof te overbruggen tussen de (neo)klassieke benadering en de institutionele benadering. Alles bij elkaar ge-

Te beginnen met de private en publieke sector, zullen bepaalde samenstellingen sectorspecifieke startups ondersteunen (Malecki, 2018). Dit wordt bijvoorbeeld bevestigd door inzichten uit de onderzoekslijn van gerelateerde industrieën (Xiao, Boschma, & Andersson, 2018). Zo zullen bedrijven eerder geneigd zijn zich in een regio te vestigen en zich er te ontwikkelen als ze gerelateerd zijn aan reeds bestaande industrieën in die regio (Neffke, Henning, & Boschma, 2011). Andere stromingen in de literatuur geven eveneens aan dat startup-formatie in de eerste plaats afhankelijk is van het aantal gevestigde bedrijven in een regio, aangezien elk bedrijf een potentiële klant of bron voor *spin-offs* vormt (Arthur, 1994; Sorenson en Audia, 2000; Stuart en Sorenson, 2003; Stel & Suddle, 2008; Li, Goetz, Partridge, & Fleming, 2016).

Spin-offs kunnen in brede zin worden gedefinieerd als nieuwe bedrijven die zijn opgericht door personen die in het jaar voorafgaand aan de oprichting werkten bij bestaande bedrijven in een aanverwante sector (Agarwal, Echambadi, Franco, & Sarka, 2004). In meer beperkte zin – maar ook meer in het algemeen – wordt met een *spin-off* verwezen naar een bedrijf dat rechtstreeks vanuit een ander (groter) bedrijf wordt opgericht. Dergelijke *spin-offs* zullen met name afkomstig zijn van de grotere

²² Volgens Brown & Mason (2017) ontbreken bij de vroege conceptualisaties van ondernemersecosystemen de sociale en culturele factoren (zoals benadrukt in de behaviourale en institutionele locatietheorieën); de tijdsdimensie (evolutionaire economische geografie); en een volledige integratie van de complexiteit van de sociaal-ruimtelijke context.

gevestigde bedrijven die fungeren als voedingsbodems voor beginnende ondernemers (Mayer, 2013b). Naast het feit dat gevestigde bedrijven *spin-offs* kunnen voortbrengen, blijkt uit studies dat *spin-offs* een systematisch hogere overlevingskans hebben in vergelijking tot andere nieuwe bedrijven (Klepper, 2002; Boschma & Wenting, 2007; Dahl & Reichstein, 2007).

Een indirect bijkomend effect van grote gevestigde bedrijven dat in een ander artikel van Mason (2013a) ter sprake wordt gebracht, is dat ze in staat zijn geschoolde arbeidsmigranten aan te trekken (Mason, Cooper, & Harrison, 2004). Deze grote gevestigde bedrijven kunnen belangrijke eerste klanten zijn voor *spin-offs* die afhankelijk zijn van arbeidsmigranten, maar ook fungeren als eerste werkgever voor arbeidsmigranten en op die manier een potentiële opstap voor een *spin-off* zijn. Voorgaande principes van grote gevestigde bedrijven als 'klant voor' of 'opstap naar' potentiële *spin-offs*, gaat niet alleen op voor externe arbeidsmigranten, maar is vanzelfsprekend ook van toepassing op (geschoolde) arbeidskrachten uit de eigen regio.

Naast *spin-off*-varianten uit het bedrijfsleven, betreft een veel voorkomende variant de *spin-offs* uit kennisinstellingen zoals universiteiten. Er kan wat betreft dit fenomeen – net als bij *spin-off*-varianten uit het bedrijfsleven – sprake zijn van zowel informele als meer formele *spin-off*-varianten. Meestal refereren *spin-offs* aan de formelere variant, waarbij nieuwe startup-activiteiten gebaseerd moeten zijn op technische ideeën die direct gegenereerd zijn in een bedrijf of kennisinstelling (Pirnay, Surlemont, & Nlemvo, 2003). Een nog striktere definitie stelt als aanvullende eis dat de overgang van het bedrijf of de kennisinstelling naar een startup direct moet zijn, zonder een tussentijdse baan ergens anders (McQueen & Wallmark, 1982). Pirnay et al. (2003) merken op dat *spin-offs* uit kennisinstellingen in veel onderzoeken slecht worden gedefinieerd, en komen tot de conclusie dat veel gehanteerde definities en dus ook de resultaten sterk verschillen. De minder formele variant van een universiteit-*spin-off* stelt dat startup activiteiten gebaseerd moeten zijn op kennis gegenereerd in een universitaire instelling, met als enige eis dat een voormalige lid of student van de onderzoeksinstelling ook daadwerkelijk zelf de (mede) oprichter van de startup moet zijn (Rappert, Webster, & David, 1999).

Kennisinstellingen in het ondernemende ecosys-

teem worden ook in het bibliometrische onderzoek van Malecki (2018) opgemerkt als belangrijke spil voor de oprichting van startups. Meer specifiek kunnen volgens Guerrero, Urbano, Fayolle, & Klofsten (2016) zogenaamde 'ondernemersuniversiteiten' beschouwd worden als belangrijke drijfveren voor startup-formatie. Andere onderzoekers, zoals Brown (2016), komen tot de conclusie dat de rol van ondernemersuniversiteiten juist wordt overschat. Brown presenteert empirisch bewijs dat suggereert dat de spillovers van universiteiten – en met name in sommige relatief afgelegen Schotse regio's – sterk worden overdreven. Een studie van Egel, Gottschalk, & Rammer (2004) concludeert tevens dat de nabijheid van de moederinstelling van weinig belang is voor locatiebeslissingen van onderzoeks-*spin-offs* uit de publieke sector; locatiebeslissingen zijn eerder afhankelijk van potentiële klanten en het aanbod aan hooggekwalificeerde arbeidskrachten. Andere studies tonen juist een positieve invloed van de kennisoutput van een universiteit in combinatie met het innovatieve vermogen van de regio, op het aantal gerelateerde startups rondom (met een gemiddelde afstand van 17 kilometer) de respectievelijke universiteiten (Audretsch & Lehmann, 2005). Niet onbelangrijk is de constatering van Guerrero, Urbano, Cunningham, & Gajon (2018) die stellen dat de beslissing om als afgestudeerde student een startup te beginnen, ook sterk beïnvloed door de arbeidsstabiliteit van het moment. Verschillende momentopnames van arbeidsstabiliteit in verschillende regio's kunnen mogelijk ook verschillen verklaren tussen afzonderlijke onderzoeken. Volgens de academische literatuur kan, ondanks verschillende bevindingen, gesteld worden dat de invloed van kennisinstellingen onomstreden is. Wel zal deze naar verwachting verschillen per regionaal ecosysteem en is er enige discussie over de exacte omvang van de effecten.

Bedrijven in de private en publieke sector genereren (in)direct *spin-off*-varianten, maar zijn ook belangrijk als potentiële klant of samenwerkingspartner. Al eerder, is aan bod gekomen dat bepaalde vormen van specialisatie en diversiteit in de private en publieke sector, innovatie en startup-formatie kunnen bevorderen. Diversiteit wordt over het algemeen gevonden in grotere steden, en daarom al snel in verband gebracht met urbanisatie-economieën. Urbanisatie-economieën zijn de voordelen die voortvloeien uit een grotere stedelijke omvang en dichtheid. Volgens Arauzo-Carod, Liviano-Solis, & Manjón-Antolín (2010) vallen urbanisatie-

Ondernemersactoren

Ondersteunende en mentorende diensten voor:

- Beginnende startup-ondernemers
 - Beginnende (vroeg stadium) ondernemers
 - Seriële en portfolio-ondernemers
- Bedrijfsincubators en co-working plekken
Netwerk- en acceleratie programma's

Leveranciers van ondernemingsbronnen

- Financiële ondersteuning (banken, risico kapitaal)
- Business angel-netwerk
- Crowd funding & peer-to-peer-diensten
- Toegang tot aandelenmarkt voor kleine bedrijven
- Relaties met grote bedrijven
- Relaties met universiteiten en R&D-centra

Ondernemende verbinders

- Professionele organisaties
- Ondernemersclubs en startup-gemeenschappen
- Zakelijke centra
- Matching-service investeerders en investeerders
- Zakelijke makelaars

Ondernemerscultuur

- Sociale status van zzp'ers, kleine bedrijven en ondernemerschap
- Rolmodellen
- Educatieve ondernemersprogramma's
- Zakelijke migratieprogramma's
- Faaltolerantie en omarming van innovatie

Figuur 2.9 Belangrijke actoren en onderlinge relaties binnen ondernemende ecosystemen. Bron: Mason & Brown (2014).

economieën onder de neo-klassieke locatietheorieën. In plaats van ruimtelijke externe effecten die worden gegenereerd door de lokale aanwezigheid van bedrijven uit sectoren, ontstaan urbanisatie-economieën door het delen van infrastructuur, instellingen of andere voordelen die verbonden zijn aan het zich bevinden in een grotere stad. Arauzo-Carod et al. (2010) analyseren verschillende empirische studies en komen tot de conclusie dat het belang van urbanisatie-economieën en de vervoersinfrastructuur verschilt per sector. Hoewel de vestigingskosten in grotere steden meestal hoger zijn, worden urbanisatie-economieën (tot op zekere hoogte) over het algemeen geassocieerd met hogere aantallen (innovatieve) startups.

De verklarende kracht van het ondernemende ecosysteem reikt verder dan rekening houden met de aanwezigheid van bedrijven, instellingen en infrastructuur. In figuur 2.9 zijn de belangrijkste actoren en onderlinge relaties die samen een gunstig ondernemersklimaat vormen overgenomen uit het artikel van Mason & Brown (2017). Ondanks het feit dat veel van deze actoren en onderlinge relaties moeilijk te kwantificeren zijn, is het belangrijk om volledig te zijn en rekening te houden met de mogelijke invloed van al deze aspecten. Dit is ten eerste nodig om de juiste aannames te kunnen doen, en vervolgens ook om de resultaten van een analyse correct te kunnen interpreteren. De rol van verschillende actoren en onderlinge relaties kan tevens worden teruggekoppeld aan paragraaf 3.1.1 waarin al werd geconcludeerd dat toegang tot zowel lokale als mondiale kennis en middelen

van aanzienlijk belang zijn voor het opstarten en ontwikkelen van een startup-bedrijf.

Een aspect van ondernemende ecosystemen dat vaak over het hoofd wordt gezien is de fundamentele rol die sociale en culturele factoren spelen bij het vormgeven van ondernemerschap (Venkataraman, 2004). Cultuur is sinds de introductie als één van de zes domeinen in figuur 2.8 nog nauwelijks besproken, maar komt voor een groot deel terug in het schema van figuur 2.9. Sommige aspecten uit figuur 2.9 zullen overigens hetzelfde zijn voor verschillende potentiële startups in een regio. Een betrouwbaar juridisch systeem, nationale ondersteunende programma's en subsidies zullen bijvoorbeeld alle startups in een regio gelijk beïnvloeden. In mindere mate gaat het voorgaande principe op voor andere aspecten, zoals bijvoorbeeld de stimulerende rol van sleutelfiguren in de vorm van business-angels, seriële ondernemers, rolmodellen, durfkapitalisten en mentoren. Vaak zijn dergelijke sleutelfiguren wél verbonden aan eerder succes in de regio, maar niet uitsluitend aan het micro-schaalniveau (Mason & Harrison, 2006; Harrison, Mason, & Robson, 2010). Kortom, voor het goed gedijen van ondernemerschap, zullen sommige culturele aspecten een (ver)gelijk(bar)e uitwerking hebben op verschillende delen van een regio. Andere sociaal-culturele aspecten verschillen mogelijk wél per locatie, maar zijn niet altijd even 'tastbaar' en daarom lastig om te betrekken in onderzoek op bedrijfsniveau.

Voor sociaal-culturele aspecten op het micro-

schaalniveau die zich wél laten uitdrukken in iets tastbaars, kan worden gekeken naar de rol van creatieve stedelijke locaties als plaatsen van innovatie. Naast hoge dichtheden en diversiteit, kunnen goed beloopbare stedelijke gebieden volgens Jacobs (1961, 1970) worden geïdentificeerd als primaire bron van innovatie. Florida (2002) heeft ook de rol van zogenaamd creatieve stedelijke omgevingen benadrukt als plaats waar een nieuwe groep van innovatieve actoren – de zogenaamde 'creatieve klasse' – tot bloei komt. Er is echter weinig empirisch bewijs voor de stelling dat in creatieve omgevingen meer innovatie startups ontstaan. Dit valt grotendeels te wijten aan een gebrek aan gegevens op microschaalniveau, maar misschien ook met een verschil tussen de woon- en werkplek van deze creatieve klasse dat niet altijd zichtbaar naar

voren komt op het schaalniveau van onderzoek. Wel tonen Florida & King (2016) op basis van het analyseren van durfkapitaal investeringen, dat startups in de software, media en entertainment sectoren vaak in de binnenstad zitten, terwijl startup investeringen in biotechnologie meer geclusterd zijn rond grote universiteiten en onderzoeksinstituten. Malizia & Motoyama (2018) laten zien dat er voor zowel stedelijke als suburbane gebieden aanzienlijke correlaties bestaan tussen plaatsgebonden levendheidsindicatoren en snelgroeiende bedrijfsconcentraties. Het is goed mogelijk dat de opkomst van innovatieve startups wordt gestimuleerd door meer creatieve of levendige omgeving, maar dat dit alleen geldt voor startups in specifieke sectoren.

3 Methode

Dit hoofdstuk behandelt de methode van dataverzameling en -verwerking voor het in beeld brengen van innovatieve startup bedrijven in de regio Noord-Nederland. *Webscraping* in combinatie met een web-based classificatie methode ligt aan de basis van dit onderzoek. Uit het theoretisch kader is naar voren gekomen dat voor een zo volledig mogelijke meting van innovatieve startups verschillende methoden nodig zijn. Dit is echter niet haalbaar voor een klein onderzoek. *Webscraping* is een opkomende methode en leent zich goed voor het in beeld brengen van de innovatieve en kleinere bedrijven. De keuze voor *webscraping* heeft als voordeel dat ook de kleinere bedrijven op een efficiënte manier gevonden kunnen worden. Daarnaast biedt het gebruik van websitegegevens in combinatie met ruimtelijke gegevens een aanvullend beeld van het innovatiepotentieel dat doorgaans niet met traditionele methoden kan worden verkregen. Voordat de methode van webscraping en de ruimtelijke analyse in meer detail wordt besproken, zal de onderzoekspopulatie en de steekproef die hieruit is getrokken worden toegelicht.

3.1 Onderzoekspopulatie en Steekproef

De regio Noord-Nederland is gekozen als onderzoeksgebied. Hieronder vallen de provincies Drenthe, Friesland en Groningen. De keuze voor dit schaalniveau was belangrijk omdat geografisch onderzoek naar innovatie op bedrijfsniveau doorgaans slechts een deel van de totale bedrijfspopulatie beslaat of beperkt is tot geografische gebieden zoals steden. Daarnaast waren de beschikbare gegevens het meest geschikt voor de analyse van Noord-Nederland, omdat voor deze regio in vergelijking tot andere Nederlandse regio's een relatief groot aantal bedrijfswebsites zijn geïndexeerd (zie tabel 3.3). In andere regio's, zoals de Randstad, is bovendien al veel onderzoek gedaan.

De onderzoekspopulatie bestaat uit alle innovatieve startups in Noord-Nederland. Het nader definiëren van de precieze onderzoekspopulatie en het vaststellen van de steekproef (de onderzoekseenheden), ging samen met de operationalisering van het begrip innovatieve startup. De basis hiervoor is gelegd in het theoretische kader waarin theorieën voor het meten van zowel innovatie als startups aan bod zijn gekomen. Op basis van deze literatuur konden criteria worden geformuleerd waar de onderzoekspopulatie en -eenheden aan moesten voldoen. De criteria waaraan een innovatieve startup

moest voldoen, betreffen de aspecten innovatief, nieuw, actief en onafhankelijk. Het was belangrijk dat hier zo goed mogelijk aan werd voldaan aangezien innovatieve startups – in vergelijking tot andere bedrijven – een unieke rol in de economie vervullen. Het in beeld brengen van de potentiële impact die de innovatieve startups hebben op de regionale ontwikkeling valt of staat bij een goede operationalisering. Hieronder wordt de operationalisering van de steekproef of onderzoekseenheden per criterium nader toegelicht, inclusief een aantal aspecten dat (nog) niet automatisch, maar wel handmatig met behulp van een aanvullende steekproef¹ kon worden gecontroleerd.

3.1.1 Topsector-expertise en innovativiteit

Of een bedrijf daadwerkelijk innovatief bezig is kan eigenlijk pas achteraf na evaluatie worden vastgesteld aangezien de gehanteerde definitie in dit onderzoek betrekking heeft op een combinatie van innovatieve activiteiten (input) en het resultaat (output) in de vorm van een succesvolle introductie van een nieuw product, dienst, productiemethode,

¹ Ter verduidelijking, het gaat bij deze steekproef niet om de onderzoekseenheden maar om een steekproef uit dé steekproef.

organisatievorm, openen van een nieuwe markt of het aanboren van nieuwe bronnen (Schumpeter & Backhaus, 2003). Specifiek geldt voor de steekproef dat is gekeken naar innovatie- en groeipotentieel uitgedrukt in aantallen startup bedrijven dat zich in betekenisvolle mate lijkt bezig te houden met innovatieve activiteiten. Het criterium innovatief slaat op de mate waarin een bedrijf activiteiten op de bedrijfswebsite beschrijft die kenmerkend zijn voor de innovatieve deelthema's.

Uit de literatuur (onder andere paragraaf 2.1.8) is gebleken dat in Nederland regelmatig de topsectoren (zie tabel 2.1) worden gebruikt om innovatie te classificeren (CBS, 2014; 2017). Dit biedt geen garantie voor daadwerkelijk succesvolle innovatie maar wel een indicatie van innovativiteit. De afbakening van innovatieve deelthema's (tabel 3.1) in dit onderzoek is gebaseerd op het topsectorenbeleid en de door StartupDelta gehanteerde thema's (Rijksoverheid, 2016; StartupDelta, 2018).

StartupDelta heeft als tiende thema ook informatietechnologie (ICT) in de onderverdeling opgenomen. Dit kan worden gezien als een afsplitsing van de topsector High-Tech Systemen en Materialen (HTSM). Volgens het topsectorenbeleid van de overheid (Rijksoverheid, 2019) behoort ICT echter niet tot de topsector HTSM. ICT wordt ook niet als aparte topsector beschouwd vanwege het *cross-sectorale* (hierna: X-sectorale)² karakter. Wel wordt binnen het topsectorenbeleid een prominente plaats gereserveerd voor ICT. Dit blijkt bijvoorbeeld

uit de samenstelling van ICT-roadmaps zoals dat ook gebruikelijk is voor de 'echte' topsectoren (Riksen, 2012). Volgens Riksen (2012) zijn ICT-doorbraken in alle topsectoren nodig. Dat argument zou ook voor de financiële sector gemaakt kunnen worden, maar de aanname is dat een belangrijk deel van innovaties binnen de financiële sector voor de rekening komt van ICT-startups. Zogenaamde *Fintech*-startups³ worden bijvoorbeeld geassocieerd met de ICT-sector. Voor dit onderzoek was het mogelijk om ICT als extra deelthema op te nemen. Samen met de negen topsectoren heeft dit geresulteerd in tien (innovatieve) deelthema's. Wanneer in het vervolg van dit onderzoek wordt verwezen naar de (deel)thema's of topsectoren, omvat dit ook ICT als officiële topsector.

Aangezien de thema's talrijk zijn en breed zijn geformuleerd (zie ook kader 3.2) werd de kans dat innovatieve activiteiten niet overlappen met minstens één deelthema als klein geacht. Tegelijkertijd zouden juist innovatieve startups in staat kunnen worden geacht om in te springen in niches van weinig-innovatieve sectoren, zonder te overlappen met een van de tien deelthema's. Toch is aangenomen dat het gros van de innovatieve startups onder ten minstens één thema zal vallen.

Met behulp van de aanvullende steekproef is nader gekeken in hoeverre de beschreven activiteiten stroken met de uitleg zoals deze op de website topsectoren.nl (Rijksoverheid, 2019) te vinden is. Specifiek is aandacht besteed aan de mate waarin intern topsector expertise aanwezig is. Hiervoor zijn drie mogelijke controle uitkomsten opgesteld:

Tabel 3.1 *Innovatieve deelthema's met afkorting.*

Volledige benaming volgens topsectoren.nl	Afkorting
Agri & Food	Agr F
Chemie	Chem
Creatieve Industrie	C Indu
Energie	Ener
HighTech Systemen & Materialen	HTSM
Informatietechnologie	ICT
Logistiek	Logi
Life Sciences & Health	LS&H
Tuinbouw & Uitgangsmaterialen	Tuinb
Water & Maritiem	Water

Bron: Rijksoverheid, 2019

- ✓ **Ja**; website-informatie suggereert dat de onderneming zelf de topsector-expertise in huis heeft.
- ≈ **Deels**; website-informatie suggereert dat de onderneming geen topsector-expertise in huis heeft, maar wel dat de bedrijfsactiviteiten raakvlakken vertonen met de topsectoren. Bijvoorbeeld een faciliterend, generiek of uitvoerend bedrijf dat afgestemd is op opdrachtgevers of klanten uit de topsectoren zonder zelf topsectorkennis te ontwikkelen.
- ✗ **Neer**; Onderneming is onterecht gevonden (*false positive*) aangezien het niet expliciet over interne topsector-expertise beschikt. Bijvoorbeeld

² Met *cross-sectoraal* of *X-sectoraal* wordt in dit onderzoek bedoeld dat activiteiten raakvlakken hebben met meerdere topsectoren of deelthema's.

³ *FinTech* is een brede overkoepelende term voor disruptieve technologieën in de financiële sector (Leong, Tan, Xiao, Tan, & Sun, 2017).

(arbeids)bemiddeling, groothandelaren en verschillende webportalen waaronder platforms voor nieuwsmedia en vacaturesites die slechts doorverwijzen naar nieuwsberichten of vacatures gerelateerd aan de topsectoren. Ook offline websites die onbereikbaar waren op het tijdstip van controle vallen onder deze uitkomst.

ren of diensten. Met de kennis van hoe lang een bedrijf actief is, kon de waarschijnlijkheid worden ingeschat of een bedrijf nog steeds als nieuw kwalificeert of niet. Een bedrijf dat niet langer nieuw is, heeft ofwel de overlevingstest in tijdseenheden doorstaan en is volwassen of het is niet langer levensvatbaar.

Het daadwerkelijk vaststellen of er wel of niet sprake is van innovatie is een haast onmogelijk taak aangezien hier zowel markt- als vaktechnische kennis voor nodig is (Garcia & Calantone, 2002). Om toch zo effectief mogelijk te kunnen valideren of er sprake is van innovatie is gekeken in hoeverre implementatie van innovatieve activiteiten wordt geclaimd op de website van de onderneming. Hiervoor zijn drie mogelijke controle uitkomsten opgesteld:

- ✓ **Ja**; het bedrijf claimt innovatief te zijn.
- ≈ **Twijfel**; gevallen die om uiteen lopende redenen twijfel zaaien over de innovativiteit van hun activiteiten. Bijvoorbeeld een bedrijf dat expliciet claimt innovatief te zijn, maar dit vervolgens ontkracht door een niet legitiem argument aan te dragen (zoals; het gebruik maken van 'de nieuwste innovaties' zonder zelf een duidelijke rol te spelen). Of een bedrijf dat impliciet de suggestie opwerpt dat het innovatief zou kunnen zijn, terwijl de kennis van de onderzoeker niet voldoende is om de innovatieve betekenis van bepaalde (meestal technische) activiteiten te kunnen duiden.
- ✗ **Onbevestigd**; Innovatie in een bedrijf kan niet worden bevestigd. Een bedrijf zal naar verwachting niet expliciet vermelden dat het niet aan innovatie doet. Mede daarom ligt het helemaal uitsluiten van innovatie niet voor de hand. Daarnaast kan een ogenschijnlijk traditioneel bedrijf zonder verdere aanwijzingen op de bedrijfswebsite toch innovatief zijn.

3.1.2 Startup-kenmerken

Naast de topsectorexpertise en eventuele innovatieclaims, zijn als aanvullende startup kenmerken de criteria 'nieuw', 'nieuwheid', 'actief' en 'onafhankelijk' gebruikt. Voor het criterium nieuw is de opstartdatum van een bedrijf gelijkgesteld met het moment van registratie. Met de opstartdatum als referentiepunt kon in tijdseenheden worden vastgesteld hoe nieuw een bedrijf is. De aanname hierbij is dat in de meeste gevallen het moment van registratie een goede graadmeter is van hoe lang een bedrijf al deelneemt aan de handel in goede-

Naast deze absolute kijk op het criterium nieuw, is ook aandacht besteed aan de vraag of een startup in relatieve zin nieuw is. Dit wordt aangeduid door het begrip nieuwheid en uitgelegd als het toevoegen van nieuwe dienst- of productiecapaciteit aan de economie. Een nieuwe registratie betekent niet altijd dat er sprake is van een nieuwe combinatie van middelen. Opsplitsingen van registraties zijn (om redenen van efficiëntie) samengevoegd om tellingen van reeds bestaande 'dienst of-productiecapaciteit' en dubbeltellingen te voorkomen. Het daadwerkelijk uitzoeken in hoeverre sprake is van nieuwheid, is een kwalitatief onderzoeksproces waarbij verschillende aspecten van volwassenheid een rol spelen.

Aangezien voor dit onderzoek indicatoren moesten worden gebruikt die een efficiënte kwantitatieve meting van de nieuwheid mogelijk maken, is op basis van de literatuur de maximale leeftijd van een bedrijf vastgesteld op 6 jaar. Dit is gedaan omdat de kans statistisch gezien klein is dat het maximale effect van startup bedrijven op regionale ontwikkeling na deze leeftijd nog bereikt gaat worden. Wanneer nieuwe bedrijven al direct volwassen kenmerken vertonen, voldoen deze niet aan de kwalificatie van een startup. Om dit soort bedrijven zoveel mogelijk te vermijden is naast een maximale absolute leeftijd ook een grens gesteld van maximaal 20 werkzame personen op het moment van meten.

Het criterium actief is al enigszins aan bod gekomen aangezien 'nieuw zijn' duidt op een bedrijf dat deelneemt aan de economie en daarmee ook actief is. Registraties geven tot op zekere hoogte aan of de onderneming actief is. Sommige geregistreerde bedrijven bestaan echter alleen op papier zonder daadwerkelijk actief te zijn. Dit is niet altijd problematisch aangezien de verwachting is dat de meeste inactieve registraties geen websites hebben – laat staan websites met tekstinhoud van innovatieve activiteiten – waardoor veel registraties automatisch zijn uitgesloten.

Onterecht kunnen sommige registraties echter nog

steeds niet zijn uitgesloten. Het nader analyseren van de bedrijfswebsites kan worden gezien als tweede meting voor het vaststellen van de activiteitenstatus. De vraag is in hoeverre de onderneming actief was op het moment van registratie en vooral in hoeverre de onderneming (nog) actief is op het moment dat de website tekst is binnengehaald. Het gebruik van bedrijfswebsites biedt de mogelijkheid om de activiteitenstatus van bedrijven te voorspellen met behulp van de technische website status. Omdat het onderhouden van bedrijfswebsites vereist dat bedrijven een aantal financiële- of arbeidsmiddelen mobiliseren, is het aannemelijk dat alleen actieve en gezonde bedrijven hun middelen daarvoor inzetten. Wanneer een bedrijf niet langer levensvatbaar is, kan worden aangenomen dat website-updates zullen ontbreken of dat de websites helemaal offline gaan (Blazquez & Domenech, 2017).

Ondanks de mogelijkheid om de technische website status automatisch te controleren, werd in de gebruikte dataset nog geen informatie over veranderingen op de websites opgeslagen. Informatie over de bedrijfswebsitesstatus kon echter worden verkregen met de Wayback Machine van The Internet Archive (Kahle, 2018). Dit is een openbaar en gratis digitaal archief van *snapshots* (momentopnames) van ongeveer 484 miljard webpagina's. Het haalt dagelijks websites binnen en slaat deze op zodat gebruikers toegang hebben om veranderingen van websites te achterhalen. Per topsector is een steekproef getrokken om onder andere de activiteitenstatus van de websites te controleren, maar ook of de inschrijfggegevens wel kloppen met de inzichten die verkregen zijn door handmatig de bron en hulpbronnen te controleren. Als aanvullende hulpmiddel is naast de Wayback Machine en de websiteinformatie zelf, indien beschikbaar ook de LinkedIn-informatie (premium proefaccount) van de onderneming en ondernemers geraadpleegd. Dit bleek een goede aanvulling op de websitegegevens. Ter validatie van de startup-kenmerken (leeftijd en omvang van het bedrijf) zijn wederom drie mogelijke controle uitkomsten gehanteerd:

✓ **Ja of onbevestigd;** gegevens van de Wayback-machine of LinkedIn-informatie zijn niet in tegenspraak met de registratiegegevens op een manier dat de omvang en leeftijd van het bedrijf niet langer aan de gestelde criteria voldoen. In de meeste gevallen was het mogelijk de inschrijfggegevens met betrekking tot de leeftijd en het aantal werknemers – te valideren. Soms

was er geen informatie te vinden over de exacte leeftijd of het aantal werknemers, met als resultaat dat de inschrijfggegevens onbevestigd bleven. In het laatste geval werd aangenomen dat de registratiegegevens correct waren.

≈ **Twijfel;** gevallen die – als gevolg van aanwijzingen afkomstig van de Wayback Machine of LinkedIn in tegenstrijd zijn met de inschrijfggegevens – om uiteen lopende redenen twijfel zaaien over de leeftijd of het aantal werknemers van het bedrijf. Bijvoorbeeld een suggestie⁴ dat het bedrijf ouder is of meer werknemers telt in vergelijking tot de aantallen gebaseerd op registratiegegevens, website-informatie of LinkedIn-informatie. Andere twijfels ontstaan als gevolg van aanwijzingen op samenwerkingsconstructies of werkzaamheden die zijn uitgesplitst in verschillende entiteiten.

✗ **Nee;** op basis van informatie op de bedrijfswebsite, de Wayback Machine of LinkedIn-informatie bleek de onderneming te groot⁵ of te oud en waren de inschrijfggegevens misleidend.

Voor de controle van de topsector-expertise, innovativiteit en startup-kenmerken kon steeds een validiteitscore van nul (✗), een (≈) of twee (✓) worden behaald. Een hogere score correspondeert met een hogere validiteit. De maximale score die behaald kon worden is zes (drie maal twee). Startups met een score van vijf of zes konden als zeer valide worden beschouwd.

3.1.3 Onafhankelijkheid en herkomst

Als laatste criterium van een innovatieve startup is binnen de steekproef nader gekeken naar de (relatieve) onafhankelijkheid. Op basis van de literatuur (paragraaf 2.2.5) is vastgesteld dat strikte onafhankelijk geen doorslaggevend criterium hoeft te zijn. Jonge kleine bedrijven die zich voordoen als 'authentiek' en op basis van de bedrijfswebsite, Wayback Machine of LinkedIn-informatie autonoom lijken te opereren, kwalificeren als relatief onafhankelijk en voldoen aan de gehanteerde startup-kenmerken.

Op basis van de literatuur (zowel paragraaf 2.2.5

⁴ Bijvoorbeeld wanneer de bedrijfservaring wordt gepromoot – los van een discussie over bedrijfservaring dat mogelijk wordt overdreven ten behoeve van klantwinning – is het niet duidelijk of om wat voor ervaring (binnen of buiten het bedrijf) het precies gaat.

⁵ Of er was overduidelijk sprake van een filiaal van een groter(e) bedrijf(sketen); niet authentiek, niet-autonoom, relatief afhankelijk, vooral uitvoerend en zelf niet vernieuwend.

m.b.t. implicaties van (on)afhankelijkheid als 3.1.2 en 3.1.3 m.b.t. gerelateerde diversiteit en *spin-offs*) is naar voren gekomen dat het toepassen van een onderscheid op basis van afhankelijkheid en herkomst wel waardevol kan zijn om de gegevens te verrijken. Om deze reden is achtergrondinformatie van de oprichter(s) van de gevonden innovatieve startups (dit zijn alleen de startups met de maximale validiteit-totaalscore van 5 of 6) achterhaald. Dit bleek goed mogelijk met behulp van de *website-snapshots* rondom het jaar van oprichting verkregen met de Wayback Machine, de actuele website-informatie van het bedrijf zelf en vooral met behulp van de LinkedIn-profielen van de individuele oprichter(s).

Op basis van de typologie van Agarwal, et al. (2004), is onderstaand onderscheid gebruikt om de verschillen in herkomst te identificeren:

- *Spin-off*; een startup waarvan de oprichter(s) in het jaar voorafgaand aan de formatie aan het werk was of waren in bestaande (gevestigde) bedrijven (A) of onderzoeksinstellingen (B) in dezelfde sector.
- *Incumbent-backed-entrant*; een startup dat direct door gevestigde bedrijven wordt ondersteund zoals dochterondernemingen, door de

moeder gesponsorde startups en *joint ventures*.

- *Diversifying entrant*; een startup dat voorheen in een andere vorm of samenstelling bestond in een andere sector voordat het als nieuwe onderneming de huidige sector betrad.
- *Onafhankelijk*; een startup (geen *spin-off*) waarvan de oprichter of oprichters geen rechtstreeks verband lijkt of lijken te hebben met de sector.

Wanneer *spin-off*-varianten uit zowel bedrijven (A) als onderzoeksinstellingen (B) apart worden meegetekend, zijn er in totaal 5 verschillende uitkomsten mogelijk voor de identificatie van de onafhankelijkheids-status en herkomst. Deze informatie kan helpen bij het verklaren en analyseren van de startup-formatie in Noord-Nederland.

Op basis van de in deze paragraaf samengestelde criteria voor topsector-expertise, innovativiteit, startup-kenmerken en het type nieuwkomer m.b.t. onafhankelijkheid, is nader vastgesteld in hoeverre de cases voldoen aan de verwachtingen. De samenstelling van de validiteit steekproef is nader beschreven in paragraaf 3.3.3. Eerst wordt in de volgende paragraaf nader ingegaan op de methode van dataverzameling.

3.2 Webscraping

Dit deelhoofdstuk gaat nader in op de toegepaste methode van *webscraping*. De wetenschappelijke context van deze methode is uitgebreid behandeld in het theoretisch kader (paragraaf 2.1.5 t/m 2.1.8). Een van de moeilijkheden voor de toepassing van *webscraping* is dat specifieke technische vaardigheden vereist zijn. Bovendien kost het samenstellen en ontwikkelen van de zoekmodellen veel tijd.

Voor dit onderzoek is daarom gebruik gemaakt van data die door middel van de zoekmodellen uit de tool 'Innovatiespotter' van het bedrijf 'Q-Modus' zijn verkregen. Deze methode van *webscraping*, is vergelijkbaar met de methode uit de studie van Gök, et al. (2015) waar eveneens gebruik is gemaakt van een classificatiemodel op basis van een zoekterm benadering. De Innovatiespotter zet

Tabel 3.2 Percentage ICT-gebruik bij kleine bedrijven naar bedrijfsgrootte, 2018.

Toegang en gebruik internet	Bedrijfsgrootte in aantal werkzame personen									
	2+	2	3-5	5-10	10-20	20-50	50-100	100-250	250-500	500+
bedrijven met website	80	68	80	86	93	94	96	97	97	99
website bevat:	(binnen bereik startup omvang)					(te groot voor startup)				
online bestellen, boeken of reserveren	28	21	30	33	37	33	35	36	38	43
beschrijvingen van producten, prijslijst	65	53	66	71	79	80	79	80	80	87
productaanpassing door klant	4	3	3	4	6	7	7	11	12	14
kantspecifieke info voor bezoeker	21	16	21	24	26	27	30	33	40	49
verwijzing naar sociale media	41	30	38	45	56	64	69	76	76	89

Bron: CBS, 2019

innovatie expertise in en maakt daarbij gebruik van informatie afkomstig uit externe bronnen en de bedrijfswebsites zelf om zoektermen te creëren en beter te maken. De zoektermen worden periodiek bijgewerkt om de snel veranderende wereld van innovatie bij te houden. Op deze manier kunnen zoveel mogelijk actuele thema's rondom innovatieve bedrijfsactiviteiten worden geïdentificeerd. De innovatieve activiteiten van bedrijven zijn zoveel mogelijk gekoppeld aan bedrijfsregistraties in het handelsregister van de Kamer van Koophandel (KvK).

Wat betreft de *webscraping* methode dient rekening te worden gehouden met het feit dat bedrijven niet altijd een website hebben en daarnaast kunnen variëren in hun motivatie voor het plaatsen van informatie op websites over hun activiteiten. Uit onderzoek van het CBS (2015) blijkt dat ruim een derde (550 duizend) van de 1,5 miljoen Nederlandse bedrijven in 2015 een website had. Dit betekent dat bijna 1 miljoen bedrijven op dat moment geen website hadden. Het overgrote deel (ruim 800 duizend) dat geen website had bestond uit bedrijven met slechts één werkzaam persoon.

Normaal gesproken beperkt periodiek onderzoek van het CBS zich tot bedrijven met minimaal 10 personen. In 2018 is echter een uitzondering gemaakt voor het jaarlijkse onderzoek naar ICT-gebruik door bedrijven, en zijn ook bedrijven met 2 tot 10 werkzame personen ondervraagd (zie tabel 3.2). De resultaten laten zien dat kleinere bedrijven over het algemeen minder vaak een website hebben. Toch is bij bedrijven vanaf 2 personen vaker wel dan niet een bedrijfswebsite met informatie over de bedrijfsactiviteiten in gebruik. Daarnaast is in het theoretisch kader al geconcludeerd dat verschillende studies het gebruik van websitegegevens voor onderzoek naar kleine innovatieve bedrijven zien als veelbelovend (o.a. Gök, et al., 2015; Youtie, et al., 2012). Kinne & Resch (2018) stellen dat bijna alle (significante) bedrijven tegenwoordig websites hebben. Aangezien het in dit onderzoek gaat om onvolgroeide opkomende bedrijven, waarvan sommige (nog) geen belang (denken te) hebben bij een openbare website, moet zeker rekening worden gehouden met het feit dat sommige startups innovatieve activiteiten uitvoeren die niet altijd zichtbaar zijn op een openbare bedrijfswebsite.

3.2.1 Procesbeschrijving

Innovatieve bedrijven worden gevonden door de Innovatiespotter op basis van een zoekalgoritme

Tabel 3.3 Geïndexeerde bedrijfswebsites per regio.

Regio	Totaal inschrijvingen	Totaal excl. SBI onbekend	Geïndexeerde websites (t.o.v. SBI bekend)	
Nederland	2801635	2079144	832392	40%
Provincie				
<u>Drenthe</u>	<u>67873</u>	<u>50763</u>	<u>27079</u>	<u>53%</u>
Flevoland	58882	45333	20176	45%
<u>Friesland</u>	<u>94975</u>	<u>71459</u>	<u>28723</u>	<u>40%</u>
Gelderland	323911	244648	102381	42%
<u>Groningen</u>	<u>79166</u>	<u>56928</u>	<u>27195</u>	<u>48%</u>
Limburg	153966	112709	43895	39%
N-Brabant	408785	311292	116320	37%
N-Holland	576202	421882	159380	38%
Overijssel	163544	124141	54814	44%
Utrecht	235252	177155	71555	40%
Zeeland	55522	42623	14863	35%
Z-Holland	583557	420211	166011	40%

Bron: Innovatiespotter, 2018*

* Bron afkomstig uit interne gegevens (niet publiek toegankelijk) van de Innovatiespotter.

in een big datasysteem. De toegepaste computertechniek die voor deze methode wordt gebruikt zal niet in detail worden besproken. De principes waarmee de gegevens zijn verzameld, zullen in algemene bewoordingen worden uitgelegd.

De Innovatiespotter verzamelt en indexeert relevante data over innovatieve activiteiten van alle bedrijven in Nederland. Dat begint met het indexeren van het handelsregister als belangrijke bron voor het ophalen van bedrijfswebsites. Niet voor elk bedrijf staat echter een website geregistreerd in het handelsregister en soms is een website gedateerd. Daarom worden ook aanvullende bronnen en zoektechnologieën ingezet om zo veel mogelijk bedrijfswebsites te vinden. Vanwege beperkte servercapaciteiten betekent dit overigens dat de dekking (geïndexeerde data en bedrijfswebsites, zie tabel 3.3) niet overal in Nederland even goed is. De dekking is afhankelijk van de bedrijfsprioriteit en heeft vooral te maken met wensen van klanten.

Nadat de bedrijfswebsites zijn verzameld is de tweede stap het onttrekken (*scrapen*) en analyseren van de tekstinhoud. Daarbij worden de bedrijfswebsites eerst *gecrawld* en worden de teksten gedownload met behulp van een geautomatiseerde *webscraper*. Vervolgens wordt de tekstinhoud van alle bedrijfswebsites geanalyseerd

en geclassificeerd op basis van een *web-based* methode waarbij een zoekmodel resulteert uit een iteratief proces: Deskresearch naar zoekcriteria levert een eerste verzameling trefwoorden en andere indicatoren inclusief weegfactoren op. Op basis van deze indicatoren wordt per thema een *query* gebouwd. Indicatoren kunnen onterecht gevonden worden, of onopgemerkt blijven. Dit worden *false positives* en *false negatives* genoemd. Door de zoekresultaten te beoordelen op basis van bestaande gegevens en een deel van de verwerkte websites handmatig te controleren, is het mogelijk om na te gaan welke criteria statistisch gezien de beste resultaten opleveren. Dit kan aanleiding zijn om stappen in het proces opnieuw te doorlopen door de samenstelling van de trefwoorden en bijhorende weging aan te passen voor een verbetering van het zoekmodel.

Het gewenste model hangt ook af van het specificeren van de zoekopdracht op basis van de exacte benodigdheden. Zo zorgt een sensitieve zoekopdracht dat weinig innovatieve bedrijven onopgemerkt blijven: het zoekalgoritme is dan sensitief genoeg om alle indicatoren te vinden. Maar het resulteert ook in een grotere kans dat een bedrijf onterecht voorkomt. Een specifieke zoekopdracht bevat weinig ongewenste bedrijven, omdat alleen sterke indicatoren worden gebruikt. In het laatste geval voldoen alle resultaten aan de zoekcriteria, maar is er een grotere kans dat innovatieve bedrijven onopgemerkt blijven. Voor dit onderzoek is gekozen voor een sensitieve zoekopdracht omdat de lange staart aan weinig relevante resultaten in het filteringsproces altijd nog kan worden uitgesloten. Wel is om technische en praktische redenen een aantal SBI-groepen en rechtsvormen op voorhand uitgesloten (zie tabel 3.4), om het aantal resultaten kleiner te maken en omdat volgens de Innovatiespotter (A. Denekamp, persoonlijke com-

municatie, 9 september 2018) in deze groepen vaak bedrijven gevonden worden die bij nadere inspectie niet relevant blijken te zijn.

3.2.2 Filteren van de inschrijfgegevens

De exports van de ruwe data zijn gemaakt op 2 augustus 2018. De ruwe data is afkomstig van 10 *query's* (zoekopdrachten) die voor ieder afzonderlijk deelthema apart uitgevoerd moesten worden. In feite gaat het bij de export om een lijst van inschrijvingen van bedrijven met een webadres en bijhorende relevantiescore. Op de relevantiescore wordt later ingegaan. De inschrijfgegevens bevatten naast de relevantiescore ook informatie zoals de geografische coördinaten, industrie-code, inschrijfdatum en het aantal werknemers. De stappen in het filteringsproces waren nog niet geautomatiseerd omdat het filteren ten behoeve van startup-bedrijven nog niet eerder was gedaan door de Innovatiespotter.

Om ervoor te zorgen dat de inschrijfgegevens overeenkomen met unieke webadressen van unieke bedrijven met unieke locaties, moest de ruwe data eerst worden opgeschoond. Bedrijven die om uiteenlopende redenen aan verschillende registraties staan gekoppeld, kunnen al snel leiden tot dubbele tellingen van identieke webadressen. Het kon ook voorkomen dat aan verschillende inschrijvingen van hetzelfde bedrijf verschillende webadressen (die soms erg op elkaar lijken, maar soms ook niet) verbonden zijn. Ten slotte corresponderen dubbele inschrijvingen regelmatig met ongelijke informatie. Deze verscheidenheid aan informatie moest zorgvuldig worden beoordeeld voordat de registraties konden worden samengevoegd of gemarkeerd om uiteindelijk alleen unieke individuele bedrijven te kunnen onderscheiden. De exacte volgorde, gegevens eerst beoordelen en vervolgens samenvoegen en verwijderen, was belangrijk; als oude registraties van hetzelfde bedrijf onmiddellijk zouden worden verwijderd, zouden nieuwere registraties van hetzelfde bedrijf onterecht geclassificeerd kunnen worden als een nieuw bedrijf.⁶

⁶ Vanwege de omvang van de ruwe data, moest het opschonen van de gegevens aanvankelijk voor elk van de 10 datasets (per deelthema) apart gebeuren. Bovendien kon pas na het filteringsproces op relevantie worden geselecteerd. In paragraaf 3.2.3 wordt nader uitgelegd waarom dit per deelthema apart moest gebeuren. Ondanks de toepassing van slimme automatiseringsprocessen bleek het opschonen van de data een intensief proces. De details kunnen niet allemaal apart worden besproken, maar de problemen hadden te maken met inconsistente structuur van de tien verschillende data-bestanden en het feit dat de gegevens oorspronkelijk vooral bedoeld

Tabel 3.4 *Uitgesloten groepen.*

Uitsluitingen naar SBI 2008-classificatie:	
47- Detailhandel	Q- Gezondheids- en welzijnzorg
I- Logies-, maaltijd- en drankverstreking	R- Cultuur
74.2- Fotografie	S- Overige dienstverlening
74.3- Vertalers en tolken	T- Huishoudens als werkgever
O- Openbaar bestuur	U- Extraterritoriale organisaties
P- Onderwijs	
Uitsluitingen naar rechtsvorm:	
Kerkgenootschap	Vereniging van Eigenaren
Vereniging	

Tabel 3.5 Filteringsproces van ruwe meervoudige inschrijvingen tot bruikbare individuele startups.

Step ↓	in vogorde van filtering:	Aantal cases (met alleen relevante percentages*) na stapsgewijs filteren per topsector									
		Agr.F	Chem	C.Indu	Ener	HTSM	ICT	Logi	LS&H	Tuinb	Water
0	Begintotaal query-export	7032	2104	20340	10152	8558	9426	8804	6256	2209	5368
1	Inschrijfdatum 2012-2018**	1819 26%	500 24%	6150 30%	2662 26%	2285 27%	2852 30%	2271 26%	1777 28%	492 22%	1322 25%
2	Geen faillissement***	1811	499	6136	2654	2278	2847	2262	1772	491	1321
3	Geen (financiële) holding	1660	448	5748	2456	2086	2644	2061	1616	454	1192
4	1 tot 20 werknemers	1027 62%	254 57%	3666 64%	1571 64%	1314 63%	1651 62%	1269 62%	1015 63%	280 62%	704 59%
5	Minimaal 2 werknemers	439 43%	101 40%	1281 35%	629 40%	551 42%	610 37%	528 42%	371 37%	127 45%	286 41%
6	Werknemers-check	437	99	1273	627	547	604	528	369	127	284
	Percentage na stap 4-6 t.o.v. stap 3	26%	22%	22%	26%	26%	23%	26%	23%	28%	24%
7	Alleen unieke inschrijflocaties	404 92%	94 95%	1196 94%	588 94%	510 93%	552 91%	486 92%	348 94%	119 94%	262 92%
8	Alleen unieke webadressen	374 93%	90 96%	1128 94%	551 94%	481 94%	517 94%	456 94%	318 91%	112 94%	243 93%
	a Percentage startups (uniek webadres) (stap 8a t.o.v. stap 0)	5,3%	4,3%	5,5%	5,4%	5,6%	5,5%	5,2%	5,1%	5,1%	4,5%
b	nevenvestigingen per uniek webadres	30	4	68	37	29	35	30	30	7	19
		0,4%	0,2%	0,3%	0,4%	0,3%	0,4%	0,3%	0,5%	0,3%	0,4%

* De percentages (tenzij anders vermeld) hebben betrekking op het verschil met de voorgaande stap: Bij stap 1 geeft dit bijvoorbeeld een indicatie van het relatieve aantal nieuwe inschrijvingen in vergelijking tot het begintotaal van stap 0. Percentages ontbreken bij stap 2 en 3 omdat deze minder relevant zijn (relatief kleine correcties zonder extreme verschillen tussen de thema's). Voor het percentage na stap 6 is (onder vermelding) het effect van stap 4, 5 en 6 ten opzichte van stap 3 samen genomen, dit geeft een indicatie van het effect van de uitsluiting van zowel te grote als te kleine werknemersaantallen op de definitieve selectie. De extra rij percentages bij stap 8a geeft een indicatie van de gefilterde individuele startbedrijven in verhouding tot de oorspronkelijke hoeveelheid geëxporteerde registraties per deelthema.

** bij meervoudige inschrijvingen van een bedrijf is de oudste inschrijving doorslaggevend

*** faillissement(saanvraag) en schuldsanering zijn uitgesloten, uitstel van betaling (surseance) is behouden

Nadat was gemarkeerd welke inschrijvingen corresponderen met identieke bedrijven en hoeveel unieke inschrijflocaties en webadressen hier aan verbonden zijn, is gekeken welke leeftijden op basis van de inschrijfdatum voldoen (stap 1 in tabel 3.5). Wanneer minstens één van de inschrijvingen verbonden aan een enkel bedrijf vóór 2012 heeft plaatsgevonden, zijn alle inschrijvingen van dat bedrijf verwijderd. Uit de tabel is af te lezen dat in de topsectoren Tuinbouw, Chemie en Water relatief veel oudere 'gevestigde' bedrijven zijn verwijderd, terwijl in de Creatieve Industrie, ICT en LS&H relatief veel nieuwe 'jonge' bedrijven aan de startup-leeftijd voldeden.

zijn voor de registratie van juridische entiteiten veroorzaakt veel complicaties die moesten worden verholpen.

Vervolgens zijn bedrijven met een faillissement of faillissementsaanvraag uitgesloten (stap 2). Daarna zijn ook (financiële) holdings (SBI-code 70102 en 6420) weggehaald (stap 3). Dit is gedaan omdat een holding wordt gebruikt als juridische kapstok om onderliggende BV's onder te hangen (KvK, 2017) en daarom – in tegenstelling tot dochterondernemingen – naar verwachting niet of nauwelijks innovatieve activiteiten zullen uitvoeren.

Hoeveel werknemers verbonden zijn aan individuele bedrijven was lastig uit de gegevens op te maken. Het aantal geregistreerde werknemers was soms gelijkmatig, en soms ongelijkmatig verdeeld over de uitsplitsingen in juridische entiteiten van individuele bedrijven. Het feit dat identieke aantallen geregistreerde werknemers op opvallende

wijze gelijk waren verspreid over verschillende inschrijvingen van hetzelfde bedrijf, heeft de suggestie van dubbeltellingen opgewekt. De problematiek rondom meervoudige inschrijvingen wordt volgens het werkgelegenheidsregister Lisa (zd) versterkt doordat ondernemers en de KvK het aantal werknemers in het handelsregister in de praktijk nauwelijks onderhouden.

Voordat dubbeltellingen van werknemersaantallen konden worden gecorrigeerd, is eerst het aantal inschrijvingen nog verder teruggebracht. Bedrijven met meer dan twintig (fulltime en parttime) geregistreerde werknemers op één van hun inschrijvingen, zijn uit de datasets verwijderd (stap 4). Gelijktijdig⁷ zijn ook alle bedrijven zonder geregistreerde werknemers of met alleen een individuele deeltijdmedewerker verwijderd. Na deze stap bevatten de gegevens nog wel individuele fulltime starters en mogelijk nog onterecht behouden resultaten op basis van misleidende gegevens die om redenen van efficiëntie pas zo laat mogelijk in het filteringsproces zijn gecorrigeerd. In de tabel is af te lezen dat wederom in de topsectoren Chemie, Tuinbouw en Water relatief veel – ditmaal niet oude maar (nieuwe) grotere en daarmee 'relatief volwassen' – bedrijven zijn verwijderd, terwijl binnen de deelthema's Energie, Creatieve Industrie, HTSM, LS&H en ICT relatief veel inschrijvingen met minder dan twintig werknemers zijn behouden.

Omdat onder de groep individuele starters relatief veel ruis van niet-innovatieve (zie Stam, 2014) en niet-ambitieuze (zie KvK, 2018)⁸ inschrijvingen werd verwacht, is deze groep in dit onderzoek buiten beschouwing gelaten (stap 5). Zoals te zien in tabel 3.5, ging dit ten kosten van een aanzienlijke hoeveelheid gegevens. Het was echter een noodzakelijke actie om de hoeveelheid te verwerken gegevens binnen de perken te houden. Na deze stap bleven alleen individuele inschrijvingen over met twee tot twintig (fulltime en parttime) geregistreerde werknemers. Vooral veel individuele starters in de Creatieve Industrie, ICT en LS&H zijn verwijderd, terwijl in de Tuinbouw en Agri & Food

relatief weinig individuele starters geïdentificeerd waren waardoor relatief veel bedrijven in deze sectoren alsnog konden worden behouden.

Als laatste correctie voor ingeschreven werknemers is gekeken bij welke bedrijven het aantal medewerkers leidde tot een twijfelgeval (stap 6). Een twijfelgeval deed zich in eerste instantie voor wanneer de som van het aantal werknemers verbonden aan de losse inschrijvingen van één bedrijf groter was dan het maximale aantal van twintig werknemers. Daarnaast zijn twijfelgevallen ook geïdentificeerd wanneer voor losse inschrijvingen van één bedrijf alleen individuele deeltijdmedewerker geregistreerd stonden. Bij twijfelgevallen is zorgvuldig beoordeeld in hoeverre het aannemelijk is dat deze alsnog voldoen aan de gestelde criteria, wat meestal resulteerde in het alsnog meenemen van de twijfelgevallen. Na deze stap was een definitieve selectie van relevante inschrijvingen het resultaat.⁹

Als laatste zijn in stap 7 en 8 de aantallen gecorrigeerd voor meervoudige inschrijvingen verbonden aan enkelvoudige (web)adressen. Als eerst is gekeken naar het aantal unieke inschrijfadressen (stap 7). Wanneer een individueel bedrijf, op basis van (nagenoeg) identieke webadressen, vestigingen op meerdere locaties (nevenvestigingen) heeft, is dit mogelijk interessant om te behouden in de data. Het bedrijf heeft immers gekozen voor de locaties van deze vestigingen.

Voordat gesproken kan worden over individuele startup bedrijven, moet echter rekening worden gehouden met mogelijke dubbeltellingen als gevolg van het behouden van nevenvestigingen. Als laatste stap is daarom gefilterd op aantal unieke webadressen per thema (stap 8). Dit hield in dat voor unieke webadressen met meerdere inschrijflocaties een hoofdvestiging vastgesteld moest worden.

Het vestigingstype – met soms ontbrekende gegevens of identieke aanduidingen – leverde echter in veel gevallen geen definitief uitsluitel over de exacte hoofdvestiging. In gevallen van twijfel konden zes verschillende controlegegevens worden geraadpleegd. Er is een poging gedaan een systematische aanpak te hanteren die zou moeten

⁷ Achteraf gezien zou het duidelijker en transparanter zijn geweest de verwijdering van bedrijven zonder geregistreerd personeel of met slechts één individuele deeltijdwerknemer afzonderlijke op te nemen. In verhouding tot de verwijdering van bedrijven met meer dan 20 werknemers was het effect van deze handeling echter verwaarloosbaar, er waren bovendien geen grote verschillen tussen de thema's opgemerkt waardoor het voor de interpretatie weinig uitmaakte.

⁸ Slechts acht procent van de zelfstandigen (zzp'ers) neemt personeel aan en groeit door (KvK, 2018).

⁹ Vanaf deze stap was de gecombineerde grootte van de tien datasets voldoende teruggebracht om alles samen te voegen (zie paragraaf 3.3.1).

Tabel 3.6 Controle t.b.v. bepaling hoofdvestiging.

Controlegegevens:	hoofdvestigingsindicator:
inschrijfdatum	oudste inschrijfdatum
geregistreerde werknemers	hoger aantal (vooral fulltime)
bedrijfsnaam	duidt niet op juridische entiteit of nevenactiviteiten, en eventueel wel op hoofdactiviteiten
rechtsvorm	
SBI-teksten	
Basisregistraties Adressen en Gebouwen (BAG)	sterkt afhankelijk van overige gegevens

leiden tot een duidelijke keuze voor één hoofdvestiging. Vanwege de rommeligheid van de ruwe inschrijfdata (ontbrekende gegevens, inconsistente en onzekerheid over de nauwkeurigheid) groeide tijdens het proces echter het besef dat juist de combinatie van alle controlegegevens de doorslag zou kunnen geven. Elk van de zes controlegegevens, en soms een combinatie, kon een sterkere of zwakkere beslissende factor zijn (tabel 3.6).

Ter illustratie: Uit de bedrijfsnaam kon soms worden afgeleid dat het waarschijnlijk om een juridische entiteit ging en in mindere mate om een actieve vestiging, maar soms was de bedrijfsnaam ook veelzeggend over het type (hoofd- of juist neven-) activiteiten verbonden aan de losse inschrijving. Activiteiten die duiden op face-to-face advies & bemiddeling zouden ten opzichten van callcenter-activiteiten bijvoorbeeld eerder worden verkozen als hoofdvestiging ongeacht een iets groter aantal geregistreerde werknemers bij de inschrijving van de callcenter-activiteiten. Echter, wanneer de inschrijving van de advies-activiteiten een latere inschrijfdatum zou hebben en door Basisregistraties Adressen en Gebouwen (BAG) aangeduid zou zijn als woonfunctie zou de inschrijvingen van de callcenter-activiteiten worden verkozen als hoofdvestiging.

Ondanks de poging om de gegevens zorgvuldig te beoordelen, bleek de data niet altijd toereikend om tot een duidelijke voorkeur voor een hoofdvestiging te komen. Willekeur speelt mede daarom ook een rol. Zelfs bij gevallen waarbij het vestigingstype wél direct uitsluitend gaf over de hoofdvestiging, bleek na inspectie van de andere bijhorende gegevens dat er nauwelijks een patroon te herkennen was met betrekking tot het onderscheid in hoofd- en nevenvestigingen.

Na stap 8a was een lijst van individuele startup

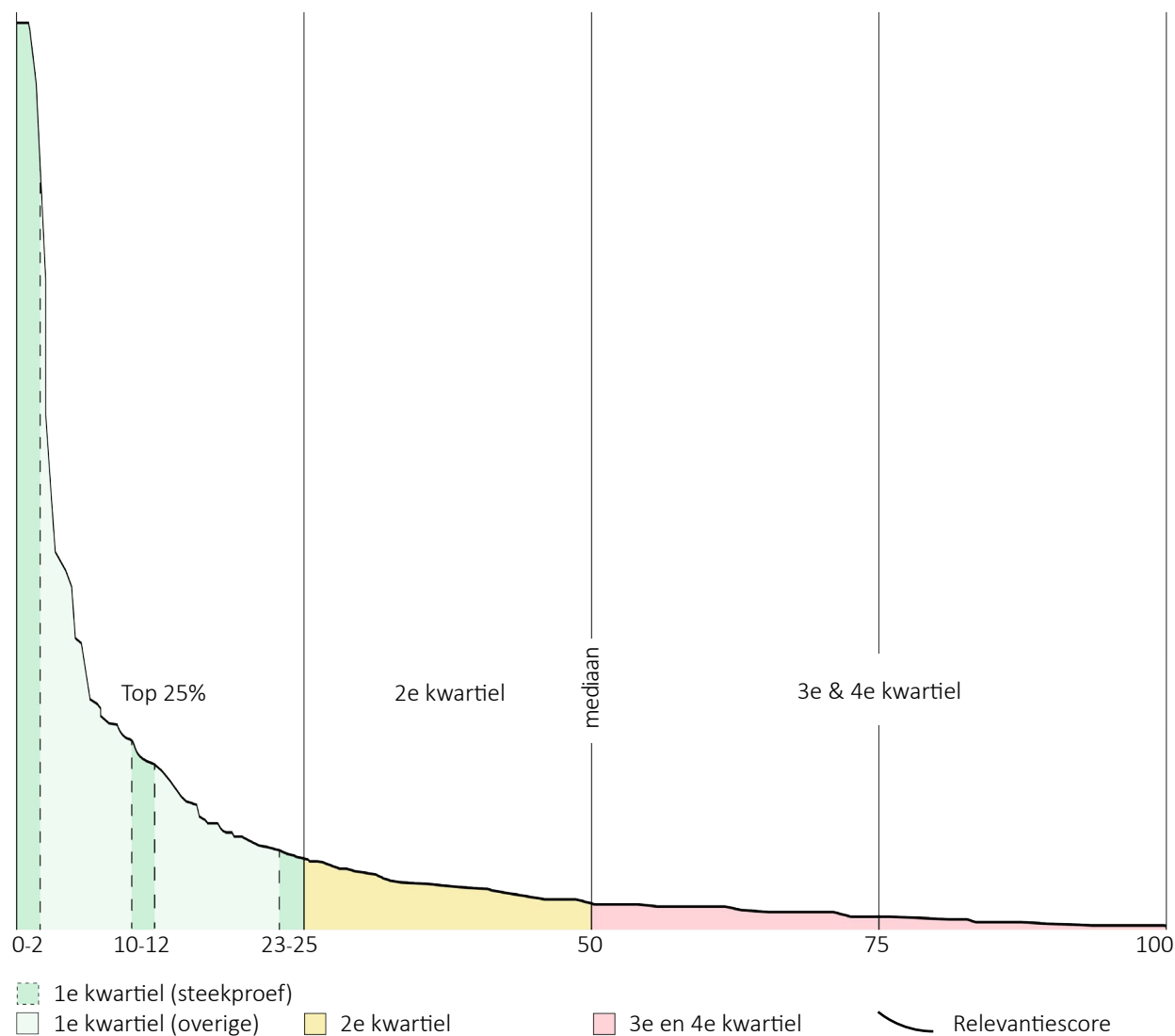
bedrijven (de hoofdvestigingen) het resultaat. De overgebleven inschrijfadressen met unieke locatie zijn gemarkeerd als nevenvestigingen (stap 8b). De aantallen na stap 8b hebben betrekking op alle nevenvestigingen, waaronder enkele uitzonderlijke gevallen van bedrijven met meer dan twee vestigingen. Uiteindelijk heeft bij elk van de deelthema's grofweg vijf tot zes procent van de totale registraties uit de export (resultaten van de *query's*) het filteringsproces (stap 0 t/m 8a) met succes doorstaan. Relatief veel inschrijvingen binnen de Chemie hebben de schifting niet overleefd. Daarentegen zijn relatief veel ICT, HTSM en Creatieve Industrie startups overgebleven.

3.2.3 Selectie op relevantiescore

Nadat het filterproces was voltooid, kwamen de resterende cases overeen met individuele bedrijven die voldeden aan de vastgestelde startup-criteria: nieuw in absolute zin (jong) en nieuw in relatieve zin (klein en dus 'onvolwassen'). Om ook te voldoen aan de operationalisering van innovatief, moesten er binnen elk van de tien deelthema's (topsectoren) selecties worden gemaakt op basis van de relevantiescore.

De startups hebben een relevantiescore gebaseerd op het aantal hits dat trefwoorden van de deelthema's hebben gemaakt met overeenkomende woorden op de bedrijfswebsites. Bepaalde trefwoorden komen vaker voor op een websites dan andere. Trefwoorden die vaker voorkomen op een website resulteren in een hogere score en komen hoger in de lijst dan bedrijven met lagere relevantiescores. De exacte scores zijn niet vergelijkbaar tussen de deelthema's en zeggen niet veel over de mate van innovativiteit of topsector-expertise. Daarentegen is de relatieve positie na rangschikking binnen hetzelfde deelthema wel van belang. Bij relatief lage scores is de statistische kans op onterechte resultaten (*false positives*) groter. In figuur 3.1 is de typische verdeling van de relevantiescore als curve verbeeld. Elk van de tien deelthema's vertoont na rangschikking op relevantiescore een piek van hoge scores die snel afloopt in een staart.

Een lange staart aan lage relevantiescores is aanwezig in iedere dataset, daarom moest een *cut-off point* (afkappunt) worden vastgesteld. Er is geen exacte lijn te bepalen waar goede resultaten overgaan in slechte. Volgens de Innovatiespotter (A. Denekamp, persoonlijke communicatie, 9 september 2018) worden de meest relevante bedrijven vaak in de eerste 10 tot 25% gevonden. Tussen de



Figuur 3.1 *Typische verdeling relevantiescores.* In aflopende volgorde van relevantiescore gerangschikt, met percentelaanduiding op de horizontale as en kwartielafstanden als *cut-off points*.

lagere scores kunnen ook relevante bedrijven zitten, maar er is ook meer ruis. Het probleem van trefwoorden is dat ze in verschillende contexten voorkomen, vandaar ook de ruis. Op basis van de ervaringen van de Innovatiespotter is gekozen het belangrijkste *cut-off point* vlak na de 1e kwartielafstand te leggen. Ter aanvulling is in de datasets als tweede *cut-off point* de mediaan gebruikt. Deze extra onderverdeling is aangebracht ten behoeve van het visualiseren, controleren en valideren van de data in een later stadium.

Een grafische weergave van de *cut-off points* is aangebracht in figuur 3.1: De *cut-off points* verdelen de data in drieën; het 1e kwartiel komt overeen met de 25% hoogste relevantiescores en wordt in het vervolg van dit onderzoek meestal aangeduid als 'Top25' of 'Top25%' en de kleur groen; het 2e kwartiel is aangeduid met de kleur geel en bevat naar verwachting meer ruis van onterecht gevonden bedrijven; in het 3e & 4e kwartiel is de lange staart met naar verwachting weinig relevante resultaten

te vinden, en aangeduid met een rode tint.

In het vervolg van dit onderzoek is voornamelijk gewerkt met de Top25 selectie – soms in vergelijking tot de rest van de data. Ook voor de steekproef is gekozen binnen de Top25 selectie te blijven omdat daar minder ruis werd verwacht waardoor een grotere hoeveelheid aan relevante resultaten gecontroleerd zou kunnen worden. Het kritisch beoordelen van de onderzoekseenheden uit een steekproef kost tijd. Daarom moest de hoeveelheid te controleren gegevens beperkt blijven. Aangezien de onderzoekseenheden verspreid waren over tien deelthema's, waarvan sommige thema's slechts een klein aantal cases bevatten, zou een aselechte steekproef mogelijk een rommelige verzameling met weinig statistische waarde opleveren. Mede daarom is besloten om gericht een meervoudige en selecte steekproef per deelthema te trekken. Voordeel van de keuze voor een beredeneerde meervoudige en selecte steekproef is dat de deelthema's op deze manier makkelijker te vergelijken

zouden zijn en verschillen eerder aan het licht zouden komen dan bij een aselechte steekproef. Per deelthema zijn het 1e tot het 2e percentiel, het 10e tot het 12e percentiel en het 23e tot het 25e percentiel in de validiteit-steekproef opgenomen (zie figuur 3.1).

Een aanvullend voordeel van het gericht steekproeven trekken uit de Top25-selectie, was dat een relatief groot deel zou overlappen met cross-sectoraal hoog scorende startups die als vierde steekproef konden worden betrokken in de steekproef analyse. Alle resultaten met drie of meer hoog scorende thema's vallen onder deze groep. Op basis van de literatuur kon worden verwacht dat veel innovaties cross-sectoraal zouden plaatsvinden omdat delen van bestaande industrieën dan worden

gecombineerd tot nieuwe innovaties. Omdat de topsectoren breed geformuleerd zijn kunnen deze onderling aan elkaar relateren. Cross-sectorale verbanden binnen individuele startups kunnen aan zowel gerelateerde (Mowery et al., 1996; Boschma & Lammarino, 2009; Boschma et al., 2014) als onge-relateerde (Enkel & Gassmann, 2010; Nootboom et al., 2007; Castaldi, et al., 2014; Janssen, 2015) diversiteit worden toegeschreven.

Door de steekproeven te richten op zowel de boven- als de ondergrens en het gebied daar tussenin, kon een goed beeld worden geschetst van de relevantie van de complete Top25-selectie als belangrijkste onderzoeksdoelgroep voor de ruimtelijke analyse.

3.3 Ruimtelijke analyse

Dit hoofdstuk bespreekt ten eerste de voorbereidingen en keuzes die gemaakt moesten worden om gegevens te selecteren en klaar te maken voor de ruimtelijke analyse. Vervolgens wordt ingegaan op de gekozen statistische methoden inclusief de aangetroffen beperkingen voor het verklaren en analyseren van startup-formatie in Noord-Nederland.

3.3.1 Samenvoegen, geocoderen en koppelen van externe data

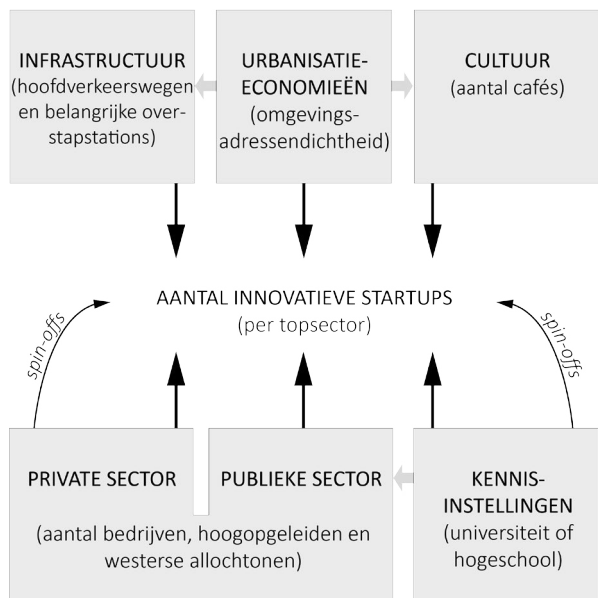
Halverwege het filteringsproces van ruwe inschrijvingen, was de gecombineerde grootte van de tien datasets voldoende teruggebracht om alles te kunnen samenvoegen. Om dubbelstellingen te voorkomen zonder daarbij de rangorde van relevantiescores of andere belangrijke inschrijfinformatie te verliezen, moest de data zorgvuldig worden gecodeerd. Het samenvoegen van de verschillende datasets ging gepaard met het creëren van nieuwe variabelen over de cross-sectorale verbanden (identieke startups met relevantiescores op meerdere thema's). In totaal waren er na het samenvoegen van de data 1586 unieke startup bedrijven (unieke webadressen) met 118 nevenvestigingen (unieke locaties).

Voor het gebruik van de geografische gegevens waren adressen en geografische coördinaten beschikbaar in de geëxporteerde datasets. Coördinaten ontbraken echter voor de topsector LS&H. Met behulp van de online tool van Morse (2005) zijn adressen van de LS&H-dataset omgezet in geografische coördinaten. Tevens konden fouten (adressen die niet overeenkwamen met de geografische

coördinaten) in de andere negen datasets met behulp van dezelfde tool worden gecorrigeerd.

Vanwege de geografisch gedetailleerde websitegegevens op micro-schaalniveau, was het mogelijk om op het kleinste schaalniveau van buurten te meten. Postcodegebieden zijn niet gekozen. Belangrijkste reden hiervoor is dat de meest relevante hulpstatistieken afkomstig zouden zijn van het CBS en al op buurtniveau beschikbaar waren. Bij gebruik van postcodegebieden zouden alle statistieken moeten worden geconverteerd met behulp van koppelta-bellen, wat mogelijk extra problemen veroorzaakt. Bovendien was de buurtindeling ten opzichten van de viercijferige postcodegebieden logischer ingedeeld in termen van sociaal-economische structuur en logische begrenzingsen, terwijl de schommelende grootte en intern minder homogene postcodegebieden tot verstoringen zou kunnen leiden.

De volgende stap was het aanmaken van buurt- en gemeente variabelen. Voor het raadplegen, bewerken en analyseren van geografisch data is het open source softwareprogramma QGIS gebruikt. Met QGIS konden geografische coördinaten van startups aan buurten worden gekoppeld. De CBS geometrie van de wijk- en buurtkaarten 2018 (CBS, 2018; Kadaster, 2018) zijn ingeladen in QGIS. Vervolgens zijn de geografische punten van de startups per geometrisch buurtoppervlak geteld en geëxporteerd. Buurtvariabelen konden eenvoudig naar een hogere schaal worden geaggregeerd om ook variabelen voor de gemeenten te creëren.



Figuur 3.2 Conceptueel model. 'Tastbare' intraregionale determinanten voor innovatieve startups, samengesteld vanuit de ecosysteem-benadering.

Op basis van de beschikbaarheid van gegevens en de bevindingen uit het theoretisch kader zijn relevante verklarende hulpvariabelen toegevoegd die in de volgende paragraaf nader zullen worden toegelicht. Aangezien de geografische indeling van 2018 is gebruikt, moesten de hulpvariabelen in sommige gevallen op deze indeling worden aangepast. In Noord-Nederland hebben de afgelopen jaren veel herindelingen voor zowel buurten als gemeente begrenzingen plaatsgevonden. Daarnaast is per hulpvariabele verschillend omgegaan met ontbrekende waarden in buurten. De gekozen werkwijze en oplossing voor zowel de herindelingen als ontbrekende waarde is nader beschreven in bijlage B2.

3.3.2 (On)afhankelijke variabelen

In figuur 3.2 en tabel 3.7 is een overzicht weergegeven van de meest relevant en beschikbaar bevonden onafhankelijke variabelen die het aantal startups in een buurt als afhankelijke variabelen zouden moeten verklaren. De selectie is samengesteld vanuit de ecosysteem-benadering en aanvullende geografische theorieën (deelhoofdstuk 2.3). De uiteindelijke keuze is vooral gebaseerd op een combinatie van geografische relevantie, beschikbaarheid van de data en de meetbaarheid in termen van effectiviteit en efficiëntie van de beschikbare determinanten. Het effect van financiële instellingen is bijvoorbeeld op intraregionaal niveau lastig te meten, bovendien is aangenomen dat deze effecten niet relevant genoeg zijn omdat dit effect voor een groot deel regionaal overschrijdend zal zijn. Daarmee is de definitieve selectie vooral gebaseerd op 'waarneembare' determinanten die op basis van theoretische gronden op intraregionaal schaalniveau verschillend tot uiting kunnen komen bij een ruimtelijke analyse.

In het theoretisch kader werd het bestaande aantal bedrijven al genoemd als belangrijkste verklaring voor nieuwe startup-formatie. De keuze voor het aantal bedrijfsvestigingen in het jaar 2016 is vooral gemaakt op basis van niet-jaarlijkse beschikbaarheid. Toch kunnen de gegevens uit het jaar 2016 worden beschouwd als een geschikte 'tussentijdse' momentopname aangezien de te verklaren startups tussen 2012 en 2018 zijn opgericht. Voor de bestaande bedrijvensvoorraad kon slechts één jaartal worden gehanteerd, daarom is aangenomen dat het aantal bedrijven per buurt in 2016 niet veel zou verschillen van de werkelijke aantallen in

Tabel 3.7 Beschrijvende statistieken van de onafhankelijke (hulp)variabelen.

Onafhankelijke variabele per buurt (tenzij anders vermeld):	N / ≠ 0	NA	min	-	max	medi-aan	gemid-delde	SE	jaar
Bedrijven in 2016	2045 / 2009	33	0	-	119	3,5	6,7	9,3	2016
OAD	2067 / 2067	11	4	-	6371	112	484,7	877,9	2018
Hoogopgeleiden (in gemeente)	2074 / 2074	4	606	-	64055	5823	10191	13931	2017
Cafés binnen 1 km	2004 / 893	74	0	-	97	0	1	7	2016
Afstand tot oprit hoofdverkeersweg (in 4 categorieën)	2078 / 2078	0	1	-	4	2	1,7	0,8	2016
Afstand tot belangrijk overstapstation (in 4 categorieën)	2078 / 2078	0	1	-	4	2	2,2	1,0	2016
Westerse allochtonen	2078 / 1648	0	0	-	1080	10	50,7	110	2018
Westerse allochtonen (wijk)	2078 / 2078	0	0	-	3850	135	362,1	573,4	2018
Nabij Kennisinstellingen Groningen	2078 / 2078	0	0	-	1	0	0,05	0,21	-
Nabij Kennisinstelling Leeuwarden	2078 / 2078	0	0	-	1	0	0,03	0,16	-

2012 (voor de oudste startups) en in 2018 (voor de jongste startups). De sectoren overheid, onderwijs en zorg zijn door het CBS niet in deze variabele opgenomen vanwege de onbetrouwbaarheid. Deze sectoren ontbreken in de bedrijvenaantallen die als verklarende variabele in deze studie gebruikt zijn. Dit geeft mogelijk een beperkt beeld van de invloed die bedrijven uit de publieke sector en onderwijs- en zorg-bedrijven uit de private sector hebben op startup-formatie.

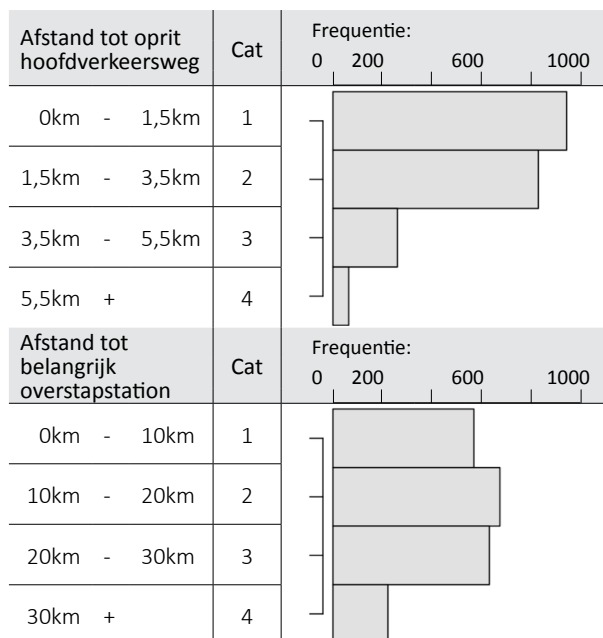
Naast de bestaande bedrijvenvoorraad, konden urbanisatie-economieën optreden als belangrijke overkoepelende verklarende variabele. Bovendien fungeerde het als een controlevariabele voor alle ontbrekende variabelen die niet afzonderlijk beschikbaar waren maar wel verband houden met urbanisatie-economieën. Er is niet gekozen om als meeteenheid van urbanisatie-economieën de bevolkingsdichtheid te gebruiken aangezien de omgevingsadressendichtheid (OAD) op het schaalniveau van de buurt als geschikter werd beoordeeld. Metingen van de OAD zijn veelomvatter en zorgen daarom op het micro-schaalniveau voor een geografisch homogener spreiding in vergelijking tot de meer heterogene spreiding van de bevolkingsdichtheid. Voor het meten van de stedelijke diversiteit is de OAD per definitie bepalender als verklarende factor voor startup-formatie, omdat deze niet alleen betrekking heeft op (homogene) woongebieden, maar ook op gebieden met diverse andere functies.¹⁰ Bedrijvenparken kunnen bijvoorbeeld worden beschouwd als onderdeel van de verklarende urbanisatie-economieën. De bevolkingsdichtheid in een klein gebied zegt echter weinig over de aanwezigheid van een bedrijvenpark, terwijl de OAD (tot op zekere hoogte) de aanwezigheid van een bedrijvenpark wél zal aangeven. Op een hoger schaalniveau zijn metingen van bevolkingsdichtheid minder problematisch omdat dan de nabijgelegen gebieden met meer woningen de heterogeniteit op het micro-schaalniveau al snel uitmiddelen. Voor verklaringen van startup-formatie in de sfeer van urbanisatie-economieën en op de kleinste schaalniveaus zoals in deze studie, is de OAD op basis van voorgaande beredeneringen het meeste geschikt bevonden als meeteenheid voor de stedelijke dichtheid van diverse menselijke activiteiten.

¹⁰ Een aanzienlijk deel van deze andere functies zal bestaan uit bedrijfsvestigingen en daarom overlappen met het aantal bedrijven per buurt in 2016. Echter bevat de OAD ook bedrijven en functies die niet bij de 'Bedrijven in 2016' zitten zoals bedrijven in de sectoren overheid, onderwijs en zorg.

Het effect van geschoolde arbeidskrachten op startup-formatie wordt doorgaans gemeten door te kijken naar de lonen of de opleiding van de beroepsbevolking. Op basis van beschikbaarheid kon alleen het aantal hoogopgeleiden in de gemeente worden gebruikt. De geografische invloed van het aantal hoogopgeleiden op de vestigingslocatie van een nieuwe startup houdt verband met de respectieve woon-werk afstand. Studies meten dit meestal op het schaalniveau van de gemeente of de arbeidsmarktregio. In het geval van specifiek startups is gekozen om op gemeente niveau te meten omdat behaviourale locatietheorieën naar verwachting een grotere rol zullen spelen dan de meer klassieke locatietheorieën (zie paragraaf 2.2.3). Met andere woorden, van ondernemers wordt verwacht dat deze in de vroege stadia van startup-formatie geneigd zullen zijn om dicht bij hun cognitieve (geografisch lastig te meten) of fysieke oorspronkelijke locatie te blijven. Cognitieve nabijheid kan ook fysieke implicaties hebben, zoals het bekend en verbonden zijn met de eigen gemeente. Fysiek nabij kan zelfs het eigen woonadres betreffen, alhoewel dit vooral zelfstandigen zullen zijn die uitgesloten zijn in dit onderzoek. Maar ondanks het uitsluiten van zelfstandigen zijn er ook flexibele woon-werk vormen denkbaar waarbij oprichters van startups zich inschrijven op het eigen woonadres. Gegevens over het opleidingsniveau waren echter niet op het micro-schaalniveau beschikbaar en zijn daarom niet nader in overweging genomen. Daarnaast zullen veel verwachte effecten het buurtniveau overstijgen. Zeker wanneer niet behaviourale locatietheorieën een grotere rol gaan spelen, bijvoorbeeld bij het opschalen van de onderneming, of bij *spin-offs* die vanuit een stabiele achtergrond opereren en daardoor veeleisender kunnen zijn aan locatie kenmerken van een nieuwe vestiging.

Wat betreft lokale cultuur en levendigheid, al dan niet geassocieerd met een creatieve gemeenschap, is het aantal cafés in een buurt gebruikt als mogelijk interessante indicator voor vestigingspatronen. Hoge aantallen cafés in een buurt kunnen – naast een suggestie voor de aanwezigheid van levendigheid of lokale creatieve gemeenschappen – ook een goede indicator zijn voor de aanwezigheid van binnensteden. Hogere aantallen startups in binnenstedelijke locaties houden mogelijk verband met de rol van creatieve stedelijke locaties als plaatsen van innovatie (Jacobs, 1961, 1970; Florida, 2002; Malizia & Motoyama, 2018).

Tabel 3.8 Herindeling bereikbaarheidsvariabelen.



De nabijheid van vervoersinfrastructuren kan worden gezien als onderdeel van de urbanisatie economieën, maar wordt vaak apart en ter aanvulling op metingen van stedelijke dichtheid in de analyse betrokken (Arauzo-Carod et al., 2010). In dit onderzoek is gekozen om de nabijheid van belangrijke overstapstations en de afstand tot een oprit van een hoofdverkeersweg in categorieën op te nemen. Dit is gedaan omdat de absolute afstanden tot beide vervoersinfrastructuren in geen enkel model een significante invloed bleken te hebben op startup-formatie, terwijl dit wel zou mogen worden verwacht. vertroebeling zou kunnen worden veroorzaakt doordat bereikbaarheid in de praktijk vooral gepaard gaat met het behalen van een bepaalde ‘drempelwaarde’ die al snel gehaald wordt omdat de infrastructuur in de meeste gebieden al snel voldoet. Kleine verschillen in absolute aantallen binnen een toelaatbare marge doen er dan weinig toe. Daarentegen kunnen grote verschillen wijzen op meer rurale of afgelegen gebieden en er wel toe doen. De herindeling van de bereikbaarheidsvariabelen en de verdeling van deze buurtkenmerken zijn in tabel 3.8 en figuur 3.2 weergegeven. De afstanden hebben betrekking op de gemiddelde afstand van alle inwoners in een buurt tot de dichtstbijzijnde oprit van een rijks- of provinciale weg (links in figuur 3.2) of het dichtstbijzijnde belangrijke overstapstation (rechts in figuur 3.2), beide afstanden zijn berekend over de weg.

In het theoretisch kader is laten zien dat arbeidsmigranten direct of indirect in verband kunnen worden gebracht met startup-formatie. In eerste instantie zullen met name de grotere bedrijven

arbeidskrachten van buiten aantrekken. Op den duur ontstaan uit deze grotere bedrijven mogelijk nieuwe *spin-offs*. Deze *spin-offs* kunnen worden opgericht vanuit de kennis van arbeidsmigranten die het gevestigde bedrijf verlaten om ideeën verder te ontwikkelen en op de markt te brengen. Dit doen zij door als oprichter of werknemer betrokken te zijn bij startup-formatie. Deze nieuwe startups – zeker wanneer het snelgroeende startups betreft – kunnen in sommige gevallen zelf ook meer arbeidsmigranten aantrekken. Specifiek, zullen westerse allochtonen worden betrokken in de analyse omdat deze vaak worden geassocieerd met arbeidsmigratie. Enerzijds heeft dit betrekking op geschoolde arbeidsmigratie bestemd voor de kennisintensieve sectoren. Anderzijds is er onder westerse migranten ook een toenemende trend zichtbaar van het aandeel lager geschoolden. Volgens het CBS en het Wetenschappelijk Onderzoek- en Documentatiecentrum (WODC) zijn dit in Nederland met name veel Polen die wel als westerse arbeidsmigrant classificeren maar geen kennismigrant zijn (Jennissen, 2011). Verder zullen niet-westerse kennismigranten ontbreken, vooral uit niet-westerse landen als India en China is bekend dat dit een grote groep kan zijn (Jennissen, 2011).

Ten slotte is voor de variabele 'Kennisinstellingen' gekozen om een dummy-variabele aan te maken voor de buurten rondom de universiteit- en de hogescholen in Groningen en Leeuwarden. Na inspectie van de resultaten (startups) bleek een mogelijk effect van deze instellingen niet verder te reiken dan 5km of de omliggende bebouwde kom van de desbetreffende gemeente. In bijlage B3.1 is de werkwijze en de selectie van de buurten rondom de kennisinstellingen nader toegelicht.

3.3.3 Ruimtelijke verklarende statistiek

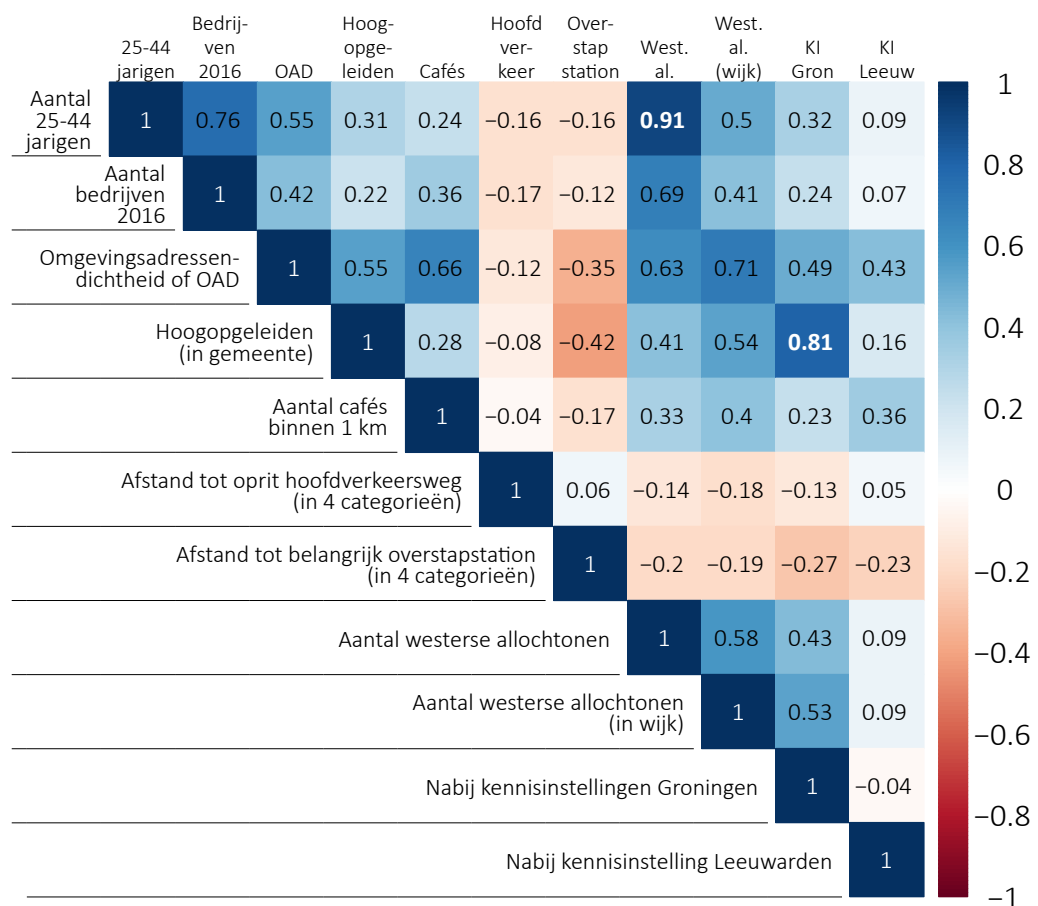
In dit onderzoek gaat de aandacht hoofdzakelijk uit naar de hoogst scorende resultaten *query*-resultaten (Top25) en de uitsplitsingen daarvan per topsector als afhankelijke variabelen van dit onderzoek. Om de hoofdvraag te kunnen beantwoorden, moesten verschillende statistische analyses worden uitgevoerd. Omdat er bij de afhankelijke variabelen sprake is van zogenaamde ‘count data’ of telgegevens zijn poisson regressies toegepast. Er was sprake van ‘count data’ omdat de uitkomstvariabele een telling is waardoor de spreiding meer beperkt is ten opzichte van een continue doorlopende numerieke variabele. Bij de poissonregressies deden verschillende problemen

zich voor die te maken hebben met het analyseren van de geografische spreiding van startups op een geaggregeerd buurtniveau. Bij de tellingen van startups per buurt moest bijvoorbeeld rekening worden gehouden met ruimtelijke autocorrelatie en te veel nullen (buurten zonder startups) in verhouding tot de resultaten, waardoor de pasvorm van statistische modellen al snel minder goed aansluit op daadwerkelijke waarnemingen.

Net als bij gewone regressies wordt bij poisson regressies uitgegaan van onafhankelijkheid van de verschillende waarnemingen. Vanuit geografische oogpunt zijn tellingen van aantallen startups per buurt niet geheel onafhankelijk. Metingen in nabijgelegen buurten kunnen bijvoorbeeld meer overeen komen (of juist niet) dan metingen in buurten die verder uit elkaar liggen. Het vaststellen van ruimtelijke patronen is belangrijk omdat gegevens die ruimtelijk gecorreleerd zijn, de analyse kunnen vertroebelen. Dit fenomeen wordt ruimtelijke autocorrelatie genoemd. Het onderliggende idee is dat het aantal startups binnen een ruimtelijk gecorreleerd gebied gedeeltelijk kan worden voorspeld op basis van informatie uit nabijgelegen buurten. Ruimtelijke autocorrelatie van waarnemingen kunnen een negatieve invloed hebben op de betrouwbaarheid van de schattingen die het model voorspelt. Voor een nauwkeuriger resultaat

moet daarom de ruimtelijke autocorrelatie worden berekend en betrokken in de analyse. Geïnspireerd door ruimtelijke econometrische technieken van onder andere Anselin (1988), Dormann, et al. (2007), Lambert, Brown, & Florax (2010) en Bivand & Wong (2018) is in bijlage B.3.1 de werkwijze toegelicht in een poging te corrigeren voor ruimtelijke autocorrelatie.

Alvorens variabelen te betrekken in de analyse zijn de onderlinge correlatiecoëfficiënten berekend (Pearson's r in figuur 3.3). Een Pearson's r coëfficiënt van +1 geeft aan dat twee variabelen perfect positief gecorreleerd zijn, terwijl een coëfficiënt van -1 duidt op perfecte negatieve onderlinge correlatie. De 25-44 jarigen als mogelijk interessant groep met oprichters correleerde sterk met het aantal bedrijven en westerse alloctonen en is niet nader in de analyse betrokken. Hoogopgeleiden in de gemeente Groningen correleerden sterk met dat van buurten rond kennisinstellingen in het stedelijk gebied van de stad. Ook de OAD correleerde matig tot sterk met verschillende variabelen. Beide variabelen zijn behouden aangezien het aandeel hoogopgeleiden ook op gebieden buiten de kennisinstellingen controleerden en de OAD op 'overige' diversiteit controleerden. De resultaten van de correlatiematrix blijven belangrijk voor de interpretatie van de ruimtelijke analyse.



Figuur 3.3 Correlatiematrix (onafhankelijke) variabelen (Pearson's r).

4 Resultaten

In dit hoofdstuk wordt eerst de kwaliteit van de gefilterde *query*-resultaten aan de hand van steekproefcontrole besproken. Vervolgens volgt een overgang van individuele cases naar resultaten van de ruimtelijke analyse waarbij de cases zijn geaggregeerd naar het buurtniveau.

4.1 Steekproefcontrole

Op basis van vooraf vastgestelde onderzoekscriteria werden de bedrijfswebsites van enkele van de best scorende gefilterde *query*-resultaten getoetst op validiteit en betrouwbaarheid. Tijdens dit proces is gebruik gemaakt van ondersteunende online informatie afkomstig van verschillende bronnen (zie paragraaf 3.1.2). De onderzoeksresultaten van de steekproef zijn gedocumenteerd in bijlage B1, gevisualiseerd in figuur 4.1 en samengevat in tabel 4.1 tot en met 4.4 en figuur 4.2 en 4.3. In totaal zijn 193 of 20% van de in totaal 939 hoogst scorende cases met een unieke bedrijfswebsite betrokken in de steekproef. 587 cases met een unieke website waren al eerder weggevallen omdat ze op geen enkel thema relatief hoog (Top25%) scoorden.

Om te beginnen geeft de visualisatie van de steekproef in figuur 4.1 – waarbij de horizontale rijen met gekleurde staven corresponderen met 254 individuele cases¹ – een goed overzicht van de data waar mee gewerkt is. De steekproefselectie bestaat uit twee hoofddelen: Boven de horizontale zwarte streep is een selectie van X-sectoraal hoogst scorende cases te vinden. Onder de streep is de gerichte samengestelde steekproef te vinden afkomstig uit 10 afzonderlijke steekproeven die zijn getrokken uit drie intervallen binnen de hoogst scorende (top25% of 1e kwartiel) *query*-resultaten per topsector (zie figuur 3.1 in het methode hoofdstuk). Een hogere ligging van de steekproefgroep in figuur 4.1 correspondeert over het algemeen met een relatief hogere relevantiescore.

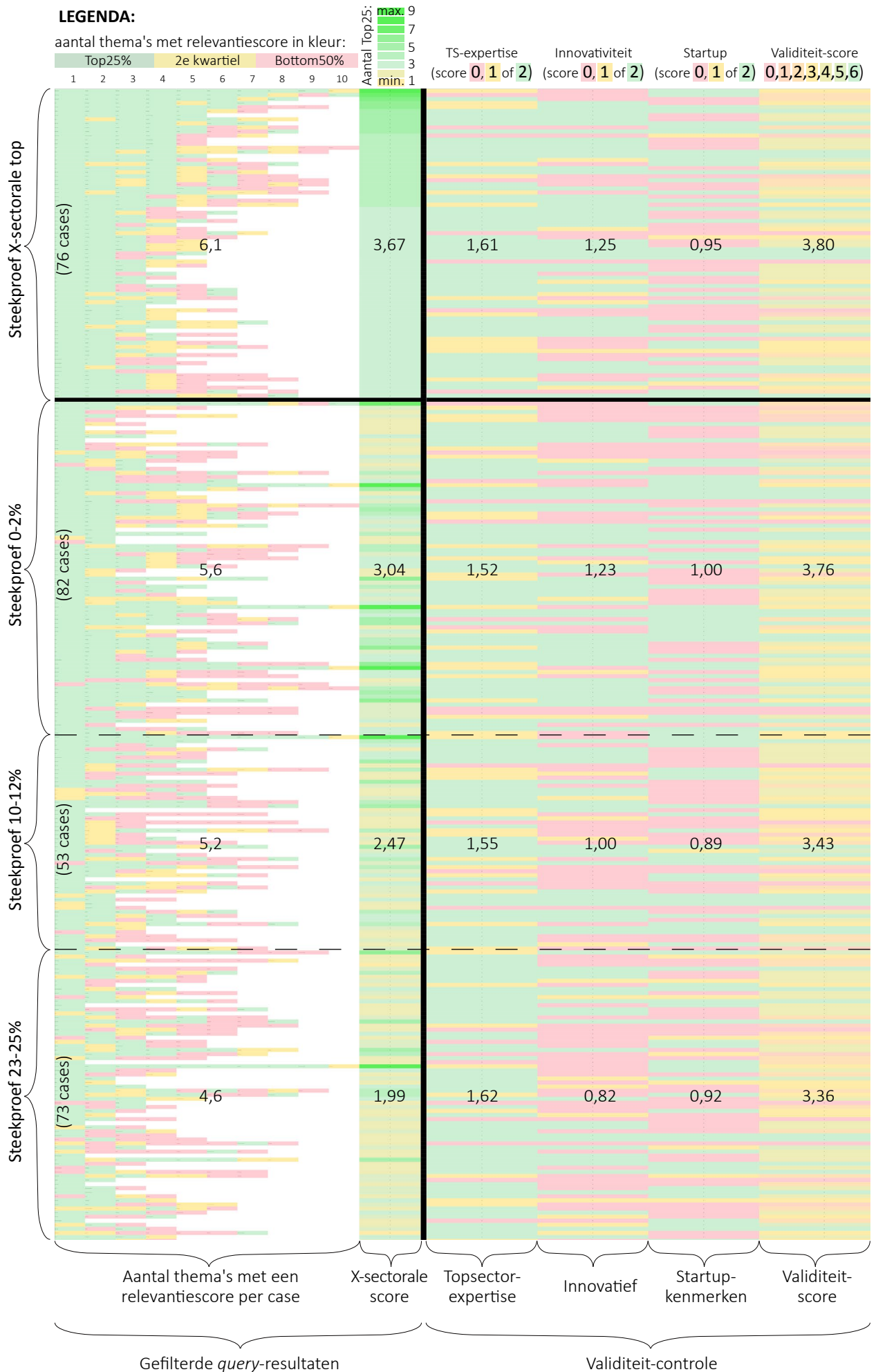
Links van de verticale zwarte streep zijn de ver-

schillende scores naar aantallen thema's te vinden: Groen correspondeert met het 1e kwartiel aan hoogste relevantiescores binnen een bepaald thema, geel met het 2e kwartiel en rood met de lange staart aan resultaten die naar verwachting minder relevant zouden zijn (het 3e en 4e kwartiel). Rechts van de streep zijn de resultaten van de handmatig uitgevoerde validiteitscontrole weergegeven.

In het midden van elke subgroep zijn in figuur 4.1 de gemiddelde scores per steekproefgroep weergegeven. De meest linkse scores hebben betrekking op het gemiddelde aantal thema's (inclusief gele en rode staven). De scores rechts daarvan, direct links van de verticale zwarte streep hebben uitsluitend betrekking op het gemiddelde aantal thema's met een relatief hoge relevantie score (groene staven uit het linker deel). Hoge scores links van de verticale zwarte streep wijzen op X-sectorale verbanden binnen de individuele bedrijven. Zoals zou kunnen worden verwacht loopt de X-sectorale score per steekproefgroep van boven naar beneden af aangezien de gemiddelde relevantiescore per steekproefgroep ook afneemt.

In de eerste kolom direct rechts van de verticale streep in figuur 4.1 is te zien dat voorspellingen voor de aanwezigheid van topsector-expertise vrij goed scoorden ongeacht de steekproefgroep. De niet valide (rode staven) resultaten hadden betrekking op *false positives* zoals bedrijven die doen aan bemiddeling (denk aan vacaturesites), nieuwsplatforms, groothandelaren, klusbedrijven en foutieve of offline websites. Bij de minder valide resultaten (gele staven) waren topsector activiteiten wel herkenbaar, maar vrij generiek of faciliterend van aard. Het gaat dan om activiteiten die globaal overlappen met veel verschillende topsectoren (denk aan

¹ Individuele cases hebben niet altijd betrekking op unieke bedrijven aangezien verschillende onderzoekseenheden uit verschillende thema's aan dezelfde bedrijven verbonden kunnen zijn. Vandaar het hogere aantal van 254 t.o.v. 194.



Figuur 4.1 Steekproef onder de startup cases. Gefilterde data onderverdeeld in subgroepen ter vergelijking van de verschillende steekproeven uit de hoogste relevantiescores (Top25%).

Tabel 4.1 Frequentietabel validiteit topsector-expertise en innovativiteit per thema.

totaal =>	Tot.uniek 193	Agr F 45	Chem 18	C Indu 95	Ener 63	HTSM 68	ICT 69	LS&H 42	Logi 46	Tuinb 9	Water 29	
Topsector- expertise score	0	26 13%	6 13%	2 11%	7 7%	14 22%	6 9%	3 4%	4 10%	7 15%	1 11%	6 21%
	1	32 17%	12 27%	5 28%	10 11%	21 33%	16 24%	9 13%	3 7%	12 26%	6 67%	9 31%
	2	135 70%	27 60%	11 61%	78 82%	28 44%	46 68%	57 83%	35 83%	27 59%	2 22%	14 48%
	gem	1,6	1,5	1,5	1,8	1,2	1,6	1,8	1,7	1,4	1,1	1,3
Innovativiteit score	0	83 43%	22 49%	8 44%	25 26%	38 60%	20 29%	20 29%	9 21%	21 46%	9 100%	16 55%
	1	23 12%	5 11%	0 0%	9 9%	6 10%	8 12%	7 10%	3 7%	3 7%	0 0%	1 3%
	2	87 45%	18 40%	10 56%	61 64%	19 30%	40 59%	42 61%	30 71%	22 48%	0 0%	12 41%
	gem	1,0	0,9	1,1	1,4	0,7	1,3	1,3	1,5	1,0	0,0	0,9

juridische expertise) of waar in ieder geval geen aanwijzingen te vinden waren voor de aanwezigheid van 'diepgaande' kennis overeenkomstig met de beschrijvingen op de websites van de topsectoren (Rijksoverheid, 2019). De groene balken in figuur 4.1 geven de meest valide resultaten weer. In het bovenste gedeelte van tabel 4.1 is de validiteit van de topsector-expertise samengevat en tevens uitgesplitst naar thema. Het totale aantal in de tabel heeft alleen betrekking op unieke cases en is exclusief dubbeltellingen van sectoroverschrijdende bedrijven afkomstig uit steekproeven van verschillende thema's. In totaal kon bij 70% van de bedrijven worden geverifieerd dat deze aan de strengste eisen voor topsector-expertise voldeden. De topsectoren Creatieve Industrie, ICT en LS&H lijken op het eerste gezicht het meest valide wat betreft daadwerkelijke ontwikkeling en aanwezigheid van topsector-kennis binnen het bedrijf. Tuinbouw scoorde daarentegen erg laag. Maar had met slechts negen cases ook het kleinste aandeel in het steekproefresultaat en daarom ook de laagste betrouwbaarheid.

De tweede kolom van de validiteit controle in figuur 4.1 en het onderste deel van tabel 4.1 tonen de toetsingsresultaten met betrekking tot innovatieclaims op de bedrijfswebsites. In de figuur is te zien dat de innovatieve claims (groene staven) per steekproef-groep toenemen naarmate de gemiddelde relevantie- en X-sectorale score per steekproef ook toenemen. Het was niet mogelijk om feitelijke innovatie te bevestigen of uit te sluiten. De score geeft enkel een indicatie op basis van claims die af te lezen zijn uit de bedrijfswebsite. In iets minder dan de helft van de gevallen bleek er na inspectie geen aanwijzing gevonden voor innovatieve activiteiten (rode staven). Sommige bedrijven verkondigden dat ze de nieuwste innovaties toepassen zonder duidelijkheid te verschaffen welke rol ze daar zelf in spelen. Dit soort gevallen

zijn als twijfelgeval gemarkeerd (gele staven). Voor bijna de helft van de cases waren wel innovatieclaims gevonden. Opvallend was dat X-sectoraal hoog scorende cases vaak ook innovatief waren met uitzondering van cases die slecht scoorde op validiteit bij de toetsing van topsector-expertise. In tabel 4.1 is verder te zien dat de sector LS&H gevolgd door de Creatieve Industrie, HTSM en ICT relatief het hoogst scoren en Tuinbouw wederom het laagst. Bij het beschouwen van de daling in validiteit ten opzichte van de topsector-expertise, valt op dat de validiteit van de Creatieve Industrie en Water sector nauwelijks lager uitvalt. Creatieve Industrie is een sector die per definitie al snel als innovatief bestempeld kan worden terwijl voor de sector Water geen directe verklaring gevonden is. De sectoren Agri & Food, LS&H, Tuinbouw, Chemie en Energie gaan relatief hard achteruit in validiteit ten opzichte van de validiteit van de topsector-expertise. Zowel figuur 4.1 als de kruistabel met unieke cases in figuur 4.2 suggereren echter een sterk verband van TS-expertise en innovativiteit. Een eenvoudige lineaire regressie is berekend om de innovativiteit te voorspellen op basis van TS-expertise. Een significante regressievergelijking werd gevonden, de voorspelde toename in innovativiteit is 0,73 per toename in TS-expertise ($\beta = 0,73$; $F(1,191) = 87,1$; $p < 0,001$, met een R^2 van 0,31).

Als derde en laatste toetsingscriteria is gekeken in

Tabel 4.2 Kruistabel validiteit TS-expertise naar Innovativiteit. (unieke cases)

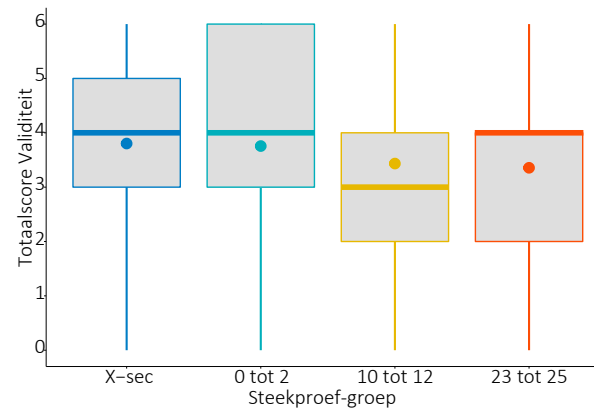
		Innovativiteit-score					
		0	1	2			
TS- exper- tise	0	26	31%	0	0%	0	0%
	1	24	29%	3	13%	5	6%
	2	33	40%	20	87%	82	94%
Totaal		83	100%	23	100%	87	100%

Tabel 4.3 Frequentietabel validiteit startup-kenmerken. (unieke cases)

Startup validiteit	absoluut	relatief
Voldoet niet, score = 0	104	54%
Twijfel, score = 1	11	6%
Ja of onbevestigd, score = 2	78	40%
Totaal	193	100%

hoeverre de cases voldoen aan startup-kenmerken. Dit heeft naar verwachting weinig te maken met de relevantiescore van de *query*-resultaten maar vooral met de validiteit en betrouwbaarheid van de inschrijfgegevens. In figuur 4.1 zijn veel rode staven te zien die wijzen op het niet voldoen aan alle criteria voor startup-kenmerken. Tabel 4.3 weergeeft alleen de unieke aantallen zonder dubbelstellingen. Ook hier is af te lezen dat ruim de helft van de cases niet voldoen aan de gestelde eisen voor startup validiteit. Sommige bedrijven waren overduidelijk niet nieuw en al volwassen. Er waren echter ook bedrijven die net buiten de vastgestelde grenzen van omvang en leeftijd vielen. De vooraf vastgestelde criteria waren echter niet als harde grens bedoeld maar moesten vooral een efficiënte kwantitatieve meting mogelijk maken. Er is tijdens het toetsen daarom soepel omgegaan met de maximale bedrijfsleeftijd van zes jaar. Wanneer er op basis van de geïnventariseerde gegevens geen duidelijke kenmerken van volwassenheid of juist aanwijzingen voor groeiambities aanwezig waren, ondanks een lichte overschrijding van maximaal 1 jaar, is oprichting in 2011 in plaats van 2012 alsnog goed gekeurd.

Dit neemt niet weg dat veel bedrijven volgens andere startup-definities wél zouden zijn goedgekeurd (zie paragraaf 2.2.3 of de verwijzing naar *Unicorns* door Stam & Saberi, 2017 in figuur 2.5) terwijl deze in dit onderzoek als niet-valide zijn geclassificeerd. Zeker wanneer bedrijven wel valide blijken op basis van topsector-expertise en innovativiteit, maar hard groeien en net meer dan twintig werknemers hebben, kan worden beargumenteerd dat ze nader zouden moeten worden overwogen. Uiteindelijk is 40% van de cases goedgekeurd op basis van startup-kenmerken. Ook dit bleef soms discutabel, zeker wanneer er geen aanvullende controle gegevens beschikbaar waren. Over het algemeen wekte het toepassen van de toetsingscriteria echter de indruk dat deze in de meeste gevallen nuttig waren om volwassen en nog niet volwassen bedrijven van elkaar te scheiden.



Figuur 4.2 Boxplots van de totale validiteitscore naar steekproef. Stippen duiden de gemiddelde score aan.

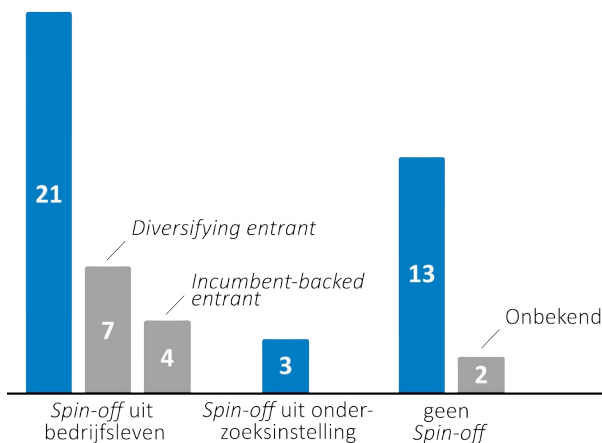
Als laatste stap, zijn de scores van de toetsingscriteria opgeteld om de totale validiteit-score vast te stellen. Deze zijn in figuur 4.1 in de laatste kolom weergegeven en onderlinge verschillen tussen de steekproef-groepen zijn in de boxplot van figuur 4.2 uitgelicht. De verschillen zijn niet heel groot maar er is wel een verband zichtbaar, namelijk dat de validiteit afneemt naarmate de steekproef-groepen relatief lagere relevantiescores bevatten. De verschillen tussen de X-sectorale hoge scores en de 0-2-steekproef (top2%) zijn erg klein. Er zijn voor zowel de drie toetsingscriteria als de totale validiteit-score eenzijdige-ANOVA toetsen uitgevoerd om de validiteit-score tussen de steekproef-groepen te vergelijken. Alleen voor het toetsingscriterium 'Innovativiteit' werd op basis van een eenzijdige-ANOVA ($F(3,280) = 3,5, p = 0,02$) een statistisch significant verschil tussen de verschillende gemiddelden van de steekproef-groepen geconstateerd. Een post-hoc-test van Tukey onthulde dat de gemiddelde Innovativiteit-score van 0,82 bij de 23-25-steekproef statistisch significant lager was dan het gemiddelde van 1,23 bij de 0-2-steekproef groep ($p = 0,03$) en het gemiddelden van 1,25 bij de X-sectorale-steekproef groep ($p = 0,03$). Er werd geen statistisch significant verschil tussen de andere steekproef-groepen aangetroffen. Bij een grotere steekproef omvang zou een mogelijk significant verschil eventueel wel geconstateerd kunnen worden. Het verschil in innovativiteit verklaard waarschijnlijk voor een groot deel de variatie in figuur 4.2.

Zowel voor innovatie- als startup-kenmerken geldt dat er sprake is van een spectrum waarin het moeilijk is een harde grens aan te houden. Een maximale validiteit-score van zes is veelzeggend, maar biedt geen garanties voor innovativiteit en groei. Een lagere score van vier voldoet niet, maar kan in veel opzichten dezelfde vergelijkbare implicaties

Tabel 4.4 Frequentietabel totale validiteit. (unieke cases)

Totale validiteit-score	absoluut	relatief
Voldoet niet, score = 0	17	9%
Voldoet niet, score = 1	13	7%
Voldoet slecht, score = 2	29	15%
Voldoet slecht, score = 3	30	16%
Voldoet gemiddeld, score = 4	54	28%
Voldoet goed, score = 5	11	6%
Voldoet goed, score = 6	39	20%
Totaal	193	100%

hebben voor regionale groei en ontwikkeling als maximale scores. In tabel 4.4 is de totale validiteit in een frequentietabel samengevat. 16% voldoet helemaal niet, 31% slecht en 28% gemiddeld. Het resterende kwart voldoet goed en heeft een hoge validiteit-score van vijf of zes. De laatstgenoemde hoogstscorende groep is nader onderzocht op (relatieve) onafhankelijkheid en herkomst.



Figuur 4.3 Verdeling van het type nieuwkomer. Op basis van de meest valide resultaten. In totaal 50 unieke startups met een validiteit-score van 5 of 6.

In figuur 4.3 is te zien dat meer dan de helft (32 van de 50) hoogst scorende startups voortkwamen uit het bedrijfsleven. Daarvan zijn vier startups aangeduid als *incumbent-backed entrants*. Deze startups worden direct ondersteund door een 'gevestigd' moederbedrijf. Bij zeven van de startups uit het bedrijfsleven waren de oprichters voorheen werkzaam in een andere sector, zogenaamde *diversifying entrants*. De meest linkse blauwe staaf in figuur 4.2 heeft betrekking op het grootse aandeel innovatieve startups en is afkomstig uit het bedrijfsleven. Samen met de diversifiërende en door een moederbedrijf ondersteunde nieuwkomers had de meerderheid van alle oprichters dus al ervaring opgedaan in dezelfde sector voordat ze de startup oprichtten.

Bij slechts drie startups was sprake van een *spin-off* vanuit onderzoeksinstellingen. Als beginselen voor de oprichting van deze startups ging het in deze gevallen om vermarkting van kennis vanuit een postdoctorale doctoraatsopleiding (Chemie), onderzoeksproject (Chemie) en onderzoeksgroep (ICT). Bij dertien van de vijftig onderzochte startups leken de oprichters geen rechtstreeks verband te hebben met de sector en leek er sprake van grote onafhankelijkheid. Opvallend bij deze groep is dat veel jonge mensen vanuit een opleiding een startup beginnen. Dit vorm een mogelijke verklaring waarom de invloed van onderzoeksinstellingen op startup-formatie niet wordt weerspiegeld door de identificatie van *research spin-offs*. Een andere mogelijke verklaring, is dat de verbondenheid aan kennisinstellingen vaak in de vroege levensloop plaatsvindt, terwijl de opgedane kennis vervolgens vooral in of met het bedrijfsleven wordt doorontwikkeld. Dat laatste is dan het meest zichtbaar.

De onafhankelijkheidsstatus wordt in de literatuur omschreven als een cruciaal aspect dat deels de unieke rol van startups verklaard. Van *spin-offs* uit onderzoeksinstellingen en startups ondersteund door moederbedrijven (*incumbent-backed entrants*) is bekend dat afhankelijkheid vaak samen gaat met ondersteuning en daarom een kracht en succesfactor kan zijn om innovaties van de grond te krijgen. Voor onafhankelijke startups is ondersteuning minder vanzelfsprekend, terwijl de ideeën en kansen van grote waarde kunnen zijn. Zoals geconcludeerd in paragraaf 2.2.5, bieden beide groepen interessante implicaties voor potentiële innovatie.

Uit de analyse van de gerichte steekproef in combinatie met de methodologische beperkingen die al gesignaleerd waren in het theoretisch kader en de methode sectie, is duidelijk geworden dat de 25% hoogste relevantiescores van de *query*-resultaten bij lange na niet representatief zullen zijn voor alle innovatieve startups in het onderzoeksgebied. In de eerste plaats is dit een gevolg van het gebruik van registratie- en websitegegevens die eigenlijk voor andere doeleinden zijn bedoeld. Wat betreft de beperkingen van websitegegevens kunnen bedrijfswebsites ontbreken, en bij wel geïndexeerde websites kon selectie of valse tellingen al snel leiden tot een vertekend beeld. Websitegegevens zijn vooral gericht op klantgerichte aspecten, en bedrijven kunnen variëren in hun motivatie voor het plaatsen (of niet plaatsen) van informatie over hun activiteiten (zie ook Van der Panne, 2007; Gök, et al., 2015). Niet alle activiteiten worden gevonden.

Inschrijfgegevens vormden vooral een beperking voor het goed inschatten van de exacte leeftijd en omvang van de startups. Dit laatste zegt iets over de mindere kwaliteit van de inschrijfgegevens die bij gebrek aan beschikbare alternatieve registers gebruikt zijn. Een ander belangrijk punt, was dat het zoekalgoritme vooral toegespitst was op het vinden van bedrijven uit de topsectoren, en niet op het vinden van innovatieve startups. Automatiseringsprocessen, die nu handmatig moesten worden uitgevoerd, hadden tot op zekere hoogte het filteren van alle *query*-resultaten mogelijk kunnen maken. Vermoedelijk had dit een optimalere validiteit-score voor startup-kenmerken als resultaat kunnen hebben.

Het feit dat de selecte steekproef beperkt re-

presentatief is voor daadwerkelijke innovatieve startups, wil zeggen dat de Top25 aan hoogste relevantiescores als afhankelijke variabelen niet geschikt zijn voor inductieve statistiek. Er kunnen op basis van de steekproef dus niet zomaar uitspraken worden gedaan over de populatie van innovatieve startups, maar slechts over de steekproefpopulatie. De steekproef-groep blijft echter een interessante groep, die belangrijke en nuttige aanvullende inzichten kan verschaffen die met traditionelere methoden niet aan het licht zouden zijn gekomen. Op basis van de steekproef-controle, is aangenaam dat de steekproefgroep sterk geassocieerd kan worden met een wat lossere definitie van een innovatieve startup. Ondanks de constatering dat lang niet alle cases daadwerkelijk als startup zullen classificeren, blijft de term startup in gebruik.

4.2 Geografische Analyse

In dit deelhoofdstuk worden de overige geografische en numerieke uitkomsten van de gegevens en tests op buurtniveau ondersteund door cijfers, tabellen en grafieken. De belangrijkste resultaten zijn geïnterpreteerd.

4.2.1 Beschrijvende statistiek

In tabel 4.5 zijn de beschrijvende statistieken van alle afhankelijke variabelen weergegeven. Uit de tabel valt af te lezen dat enkele startups niet zijn meegenomen in de ruimtelijke analyse omdat er hulpvariabelen ontbreken in 84 buurten (in tabel 3.7 waren dit er nog 2078 t.o.v. 1994 in tabel 4.5). In totaal bevinden acht Top25 startups zich in een van deze buurten. Een aantal sectoren bevat veel minder startups, dit komt overeen met de verschillen geconstateerd bij de ruwe inschrijfgegevens (zie methode sectie tabel 3.1).

De 'X'-variabelen hebben betrekking op cross-sectorale verbanden binnen de startups. 'X-T25 Som' zijn de Top25 startups vermenigvuldiging met het aantal Top25thema's binnen individuele startups. 'X-T25xT' is berekend door de Top25 startups te vermenigvuldigen met het totaal aantal thema's (inclusief lager scorende relevantiescores). 'X-T25 >1' zijn alle startups met minstens twee hoge scores (binnen top25 percentiel). 'X-T25 >1 Som' betreft dezelfde startups, maar is ditmaal gewogen door een vermenigvuldiging van het aantal Top25-thema's. Omdat een hoge mate van sectoroverschrijdende kennis geassocieerd wordt met een verhoogde kans op (disruptieve) innovatie, is het interessant om met behulp van deze variabelen te

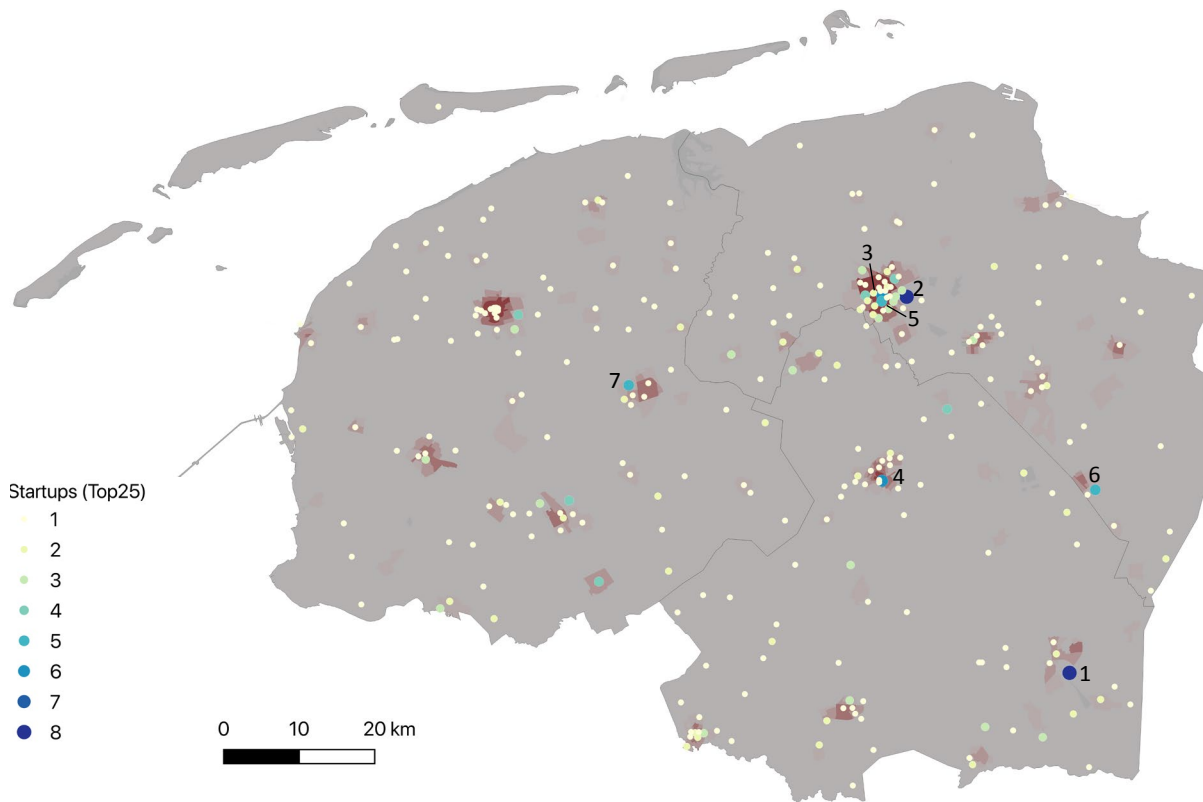
controleren op mogelijk interessante afwijkingen.

Als laatste in tabel 4.5 zijn de steekproefresultaten opgenomen. 'Ste.pr' betreft alle startups die in de

Tabel 4.5 Beschrijvende statistieken van de afhankelijke variabelen.

	N / ≠ 0	Som	X*	min - max	medi-aan	gem
Totaal	1994 / 685	1562	24	0 - 35	0	0,78
Top25	1994 / 285	413	8	0 - 8	0	0,21
Agr F	1994 / 77	81	2	0 - 2	0	0,04
Chem	1994 / 15	17	1	0 - 3	0	0,01
C Indu	1994 / 131	184	3	0 - 5	0	0,09
Ener	1994 / 94	114	1	0 - 3	0	0,06
HTSM	1994 / 72	90	1	0 - 3	0	0,05
ICT	1994 / 97	111	4	0 - 4	0	0,06
LS&H	1994 / 77	87	0	0 - 4	0	0,04
Logi	1994 / 48	53	1	0 - 3	0	0,03
Tuinb	1994 / 15	15	0	0 - 1	0	0,01
Water	1994 / 51	56	0	0 - 2	0	0,03
X-T25 Som	1994 / 285	832	-	0 - 16	0	0,42
X-T25xT	1994 / 194	926	-	0 - 19	0	0,46
X-T25 >1	1994 / 132	185	1	0 - 4	0	0,09
X-T25 >1 Som	1994 / 132	496	-	0 - 17	0	0,24
Ste.pr	1994 / 148	191	2	0 - 4	0	0,10
TOP	1994 / 46	50	0	0 - 3	0	0,03
SEMI	1994 / 76	88	1	0 - 3	0	0,04
FLOP	1994 / 51	53	1	0 - 2	0	0,03

* Deze kolom bevat het aantal startups in buurten die zijn uitgesloten in de ruimtelijke analyse omdat enkele onafhankelijke variabelen ontbreken (dit is niet apart berekend voor gewogen 'X'-variabelen').



Figuur 4.4 Unieke Top25-query-resultaten geplot op het onderzoeksgebied. Op de achtergrond de stedelijkheidsgraad (zie legenda figuur 1.1). Bronnen: CBS (2018) en het Kadaster (2018).

steekproef-controle terecht waren gekomen (zie tabel 4.4). 'TOP' heeft betrekking op de innovatieve startups met een goede validiteit score van vijf of zes, 'SEMI' betreft startups met een gemiddelde validiteit score van vier. En 'FLOP' betreft alle startups met een lagere validiteit score. Ook voor deze variabelen geldt dat kan worden gecontroleerd op afwijkingen.

Wat opvalt aan tabel 4.5 is dat er in de meeste buurten helemaal geen startups te vinden zijn. Figuur 4.4 weergeeft het onderzoeksgebied met daarop alle Top25 startups per buurt. Op deze kaart is elke buurt met minimaal één startup aangeduid als een stip. Een grotere stip geeft aan dat de desbetreffende buurt een hoger aantal startups bevat. Zoals kon worden verwacht zijn in de stedelijke gebieden relatief veel startups te vinden. En in

de landelijke gebieden weinig. Een aantal buurten in steden bevat vijf of meer startups. Deze buurten zijn in figuur 4.4 en tabel 4.6 genummerd en met behulp van allecijfers.nl (open data op basis van officiële bronnen zoals het CBS) met het type buurt gelabeld. De statistieken tonen aan dat met name buurten die worden getypeerd als bedrijventerreinen, maar ook buurten die worden getypeerd als binnensteden, de hoogste aantallen startups bevatten.

4.2.1 Verklarende statistiek

Voor de ruimtelijke analyses zijn poisson regressies uitgevoerd voor alle *query*-resultaten en de Top25-startups inclusief uitsplitsingen naar topsector.

Aangezien de variantie bij de poissonmodellen groter was dan de gemiddelden, was er sprake van 'overdispersion' (overdispersie). Overdispersie is de aanwezigheid van grotere statistische spreiding in de *query*-resultaten dan zou worden verwacht op basis van de poissonmodellen. In bijlage B3 figuur B3.3 is de verdeling van de Top25-selectie geplot met daarnaast in figuur B3.4 een rootogram (Kleiber & Zeileis, 2016) dat de pasvorm van het Top25-poissonmodel visualiseert. Ondanks over- en onderdispersie bleken de poissonmodellen zich niet slechter aan de *query*-resultaten aan te passen dan alternatieve nul-inflatie-modellen aangezien Vuong toetsing geen significante verschillen con-

Tabel 4.6 Buurten met meeste Top25 startups.

#	Buurtnaam	Plaats	Type buurt	T25-startups
1	Industrieterrein Bargerveer	Emmen	Bedrijventerrein	8
2	Eemspoort	Groningen	Bedrijventerrein	8
3	Binnenstad-Zuid	Groningen	Binnenstad	6
4	Brinkkwartier	Assen	Binnenstad	6
5	Rivierenbuurt	Groningen	Binnenstad	5
6	Industriegebied Vleddermond	Stadska-naal	Bedrijventerrein	5
7	De Haven	Smallinger-land	Bedrijventerrein	5

stateerden. Om te corrigeren voor overdispersie zijn quasi-poisson regressies (Zeileis, Kleiber, & Jackman, 2008) uitgevoerd met hogere standaardafwijkingen en aangepaste p-waarden als resultaat.

Ondanks de ontwikkelingen in de wetenschap rondom ruimtelijke econometrie (bijvoorbeeld Anselin, 1988; Dormann et al., 2007; Lambert et al., 2010; Bivand & Wong, 2018) bleek zonder uitgebreide statistische kennis geen eenvoudige oplossing voor de bepaling en uitvoering van een geschikte techniek om ruimtelijke effecten in de modellen op te nemen (zie bijlage B4). Volgens de Global Moran toetsing naar verschillende ruimtelijke conceptualisaties (zie tabel B.4.1 in bijlage B4) leek er echter weinig sprake van ruimtelijke autocorrelatie – in ieder geval niet voor de totale gefilterde dataset en de Top 25-subselectie. Het niet kunnen vaststellen van significante ruimtelijke autocorrelatie zou te maken kunnen hebben met buurten die op basis van homogene sociaal-economische kenmerken met logisch gevormde begrenzingen mogelijk meer overeenkomen met de voetafdruk van ruimtelijke processen dan wordt verondersteld in studies die intern-heterogener gebieden onderzoeken (Bivand & Wong, 2018). Ook zou het te maken kunnen hebben met de poisson modellen of de ruimtelijke gewichtenstructuren die nog niet voldoende gespecificeerd² kunnen zijn. Beperkte beschikbaarheid van onafhankelijke variabelen en tijd om econometrische kennis te absorberen en toe te passen, kan worden gezien als een beperking inherent aan de scope van deze studie.

De quasi-poissonmodellen in tabel 4.7 zijn het meest geschikt bevonden voor de ruimtelijke analyse. Om de interpretatie te vergemakkelijken en thematische verschillen bloot te leggen, zijn alleen de getransformeerde incidentie-ratio's (IRR's) met significantieniveaus weergegeven. De regressiecoëfficiënten kunnen als volgt worden afgelezen: een toename van de bestaande bedrijfspopulatie in

² Buczkowska & De Lapparent (2014) verwijzen naar studies van Partridge, Boarnet, Brakman, & Ottaviano (2012), LeSage & Pace (2014) en Vega & Elhorst (2013) waaruit blijkt dat de exacte specificatie van een ruimtelijke gewichtsmatrix een controversieel aspect van ruimtelijke methoden kan zijn. Kinne & Resch (2017) illustreren met geavanceerde geocoderingsprocessen een van de vele mogelijkheden om de ruimte ten behoeve van economisch geografisch onderzoek te conceptualiseren en modelleren. Het voorgaande bevestigt de overtuiging dat er vele mogelijkheden zijn die ongetwijfeld leiden tot betere oplossingen, maar ook dat deze oplossingen niet altijd binnen handbereik zijn.

een buurt met 1 eenheid (gelijk aan 10 bedrijven) wordt geassocieerd met een 1,06 (+6%) maal groter aantal startups in de buurt.

Bij het beoordelen van de resultaten in tabel 4.7 blijkt ten eerste dat de modellen die het totaal aan *query*-resultaten bevatten de variantie veel beter verklaren dan de modellen die enkel startups bevatten met minimaal 1 geïdentificeerd thema scorend in het 1e tot en met het 25e percentiel van de gerangschikte relevantiescores. De verklaring is dat een grotere hoeveelheid data een betrouwbaarder model oplevert waardoor uitsplitsingen met minder tellingen zwakkere modellen opleveren. De eerste kolommen met alle unieke *query*-resultaten en de top25-resultaten als afhankelijke variabelen kunnen dienen als referentiekader voor de overige modellen. De Top25-resultaten kunnen worden beschouwd als 'innovatiever' dan het totaal aan *query*-resultaten³. Tenzij anders vermeld, wordt bij de nadere interpretatie uitsluitend gekeken naar de significante resultaten.

De belangrijkste significante voorspeller is zoals verwacht het aantal bestaande bedrijven in de buurt. Onder deze al aanwezige bedrijven zitten ook startups die al voor de telling in 2016 zijn opgericht. Daarnaast valt op dat aantallen westerse allochtonen in de buurt en wijk significante voorspellers zijn. Net als aantallen cafés en bedrijven geldt voor deze variabelen dat ze matig tot sterk correleren (figuur 3.3) met de OAD. De significantie van de OAD als voorspeller valt daarom lager uit dan op basis van de theorie verwacht zou mogen worden. Voor de interpretatie van deze afzonderlijke variabelen moet daarom enige rekening worden gehouden met onderlinge correlatie.

4.2.1a Vergelijking op innovativiteit

Bij het referentie-model met alle *query*-resultaten zijn alle voorspellers significant. Wanneer dit met de significante resultaten van de naar verwachting innovatievere Top25-selectie (deels overlap) wordt vergeleken, vallen enkele kleine verschillen op:

Bij een toename van tien cafés in de buurt neemt het aantal innovatievere startups minder snel af (-21%, $p < 0,001$) dan het aantal minder innovatieve startups (-25%, $p < 0,001$).

³ Deze verwachting is gebaseerd op de theoretische en empirische bevindingen beschreven in methode sectie 3.2.3 en de ANOVA-toets in paragraaf 4.1 bij figuur 4.2.

Tabel 4.7 Quasi-poissonmodellen inclusief uitsplitsingen naar Top25 thema's.

	ALLES	Top 25	Agr F	Chem	C Indu	Ener	HTSM	ICT	LS&H	Logi	Tuinb	Water	
Voorspellers (in buurt, tenzij anders vermeld)	IRR	IRR	IRR	IRR	IRR	IRR	IRR	IRR	IRR	IRR	IRR	IRR	
(Intercept)	0,58 ***	0,15 ***	0,05 ***	0,01 ***	0,06 ***	0,02 ***	0,03 ***	0,03 ***	0,03 ***	0,03 ***	0,01 ***	0,02 ***	
Bedrijfsinschrijvingen 2016 (x10)	1,06 ***	1,06 ***	1,06 ***	1,07 ***	1,06 ***	1,07 ***	1,07 ***	1,06 ***	1,08 ***	1,05 ***	1,04 •	1,07 ***	
Omgevingsadressendichtheid (x1.000)	1,14 **	1,12	0,88	0,80	1,30 **	0,94	0,89	1,11	0,89	0,99	0,15 •	0,83	
Aantal hoogopgeleiden (in gemeente x100.000)	0,31 **	0,46	0,05 •	1,10	0,45	0,29	0,53	3,30	0,34	0,48	0,53	1,25	
Aantal cafés binnen 1 km (x10)	0,75 ***	0,79 ***	1,08	0,94	0,74 ***	0,65 *	0,78 **	0,81 **	0,83	0,82	2,49	0,82	
Afstand tot oprit hoofdverkeersweg (dummy's in 4 categorieën)	1,2	0,72 ***	0,80 •	0,56 **	0,56	0,88	0,91	0,50 **	0,92	0,62 •	0,54 **	0,95	0,73
	2,3	0,65 **	0,80	0,88	0,00	0,72	0,86	0,74	0,58	0,23 **	0,35	1,08	0,18 •
	3,4	0,53 •	0,32	0,00	0,00	0,00	0,00	0,00	0,92	0,79	0,00	0,00	0,22
Afstand tot belangrijk overstapstation (dummy's in 4 categorieën)	1,2	0,76 **	0,82	0,95	0,69	0,75	1,38	0,65	0,80	0,91	0,62	1,54	1,16
	2,3	0,76 **	0,75	0,74	0,84	0,57 **	1,73 •	0,71	0,75	0,84	0,61	0,33	0,48
	3,4	0,65 **	0,54 *	0,45	0,00	0,39 •	1,26	0,35	0,39	0,47	0,00	0,69	0,42
Westerse allochtonen (x100)	0,77 ***	0,73 ***	0,58 ***	0,64 **	0,78 ***	0,81 *	0,67 ***	0,71 ***	0,62 ***	0,80 **	0,99	0,76 **	
Westerse allochtonen (in wijk x1.000)	1,36 ***	1,35 *	1,19	1,47	1,51 *	1,57 **	2,22 ***	1,51 **	1,71 *	1,60 •	3,21 **	1,55	
Groningen KI (dummy nabijheid)	0,1	3,10 ***	3,54 *	9,67 *	6,41	1,75	5,49 **	3,69	1,99	3,24	4,52	0,18	2,27
Leeuwarden KI (dummy nabijheid)	0,1	1,99 ***	1,92 **	4,12 **	8,03 **	2,50 **	4,07 **	4,69 *	2,14 •	3,99 **	3,57 **	0,00	0,84
Observaties N / ≠ 0	1994 / 685	1994 / 285	1994 / 77	1994 / 15	1994 / 131	1994 / 94	1994 / 72	1994 / 97	1994 / 77	1994 / 48	1994 / 15	1994 / 51	
Cox & Snell's R2 / Nagelkerke's R2	0,61 / 0,69	0,19 / 0,33	0,04 / 0,16	0,02 / 0,25	0,13 / 0,33	0,07 / 0,23	0,08 / 0,31	0,08 / 0,29	0,08 / 0,31	0,04 / 0,22	0,01 / 0,14	0,05 / 0,26	

• sig<0,1; *sig<0,05; **sig<0,01; ***sig<0,001

Een verschil van vergelijkbare omvang is geconstateerd voor het aantal westerse allochtonen in de buurt. Bij een toename van 100 westerse allochtonen in de buurt, neemt het aantal innovatieve startups relatief meer af (-27%, <0,001) in vergelijking tot de afname van minder innovatieve startups uit de referentiegroep (-23%, <0,001).

Een groter verschil is geconstateerd bij de afstand tot een belangrijk overstapstation. Wanneer de afstand over de weg tot een belangrijk overstapstation groter dan 30 kilometer wordt ten opzichte van een afstand tussen de 20 en 30 kilometer, neemt

het aantal innovatieve startups beduidend meer af (-46%, p=0,046) ten opzichte van bedrijven binnen het totaal aan query-resultaten (-35%, p<0,001).

Als laatste verschil tussen de Top25 en het totaal aan resultaten, blijkt dat in buurten nabij de kennisinstellingen van Groningen de statistische kans op innovatieve startups ruim 3,5 keer groter wordt (+354%, p=0,005) terwijl dit bij de minder innovatieve referentiegroep nog maar 3,1 keer groter wordt (+310%, p<0,001). Voor kennisinstelling in de buurten van Leeuwarden is het verschil kleiner en voorspelt het model het omgekeerde. Het aan-

tal startups uit de referentiegroep zal binnen de bebouwde kom in Leeuwarden nagenoeg verdubbelen (+199%, $p < 0,001$) terwijl dit voor innovatievere startups iets minder dan een verdubbeling zal zijn (+192%, $p = 0,036$).

Samenvattend lijkt het erop dat innovatievere startups eerder in meer levendige gebieden (meer cafés in de omgeving) en dichterbij belangrijke treinstations (niet meer dan 30 kilometer) te vinden zullen zijn dan minder innovatieve startups. Ook zijn innovatievere startups relatief iets minder snel te vinden in buurten waar meer westerse allochtonen wonen. Dit heeft mogelijk ook te maken met bedrijventerreinen waar minder woningen te vinden zijn. Innovatievere startups lijken de universiteitsstad Groningen te verkiezen boven Leeuwarden dat als onderwijskennisinstelling alleen een hogeschool heeft. Andere aantrekkingskrachten van Groningen houden mogelijk verband met urbanisatie-economieën of netwerk-mogelijkheden.

4.2.1b Vergelijking op thema's

Een belangrijk doel van dit onderzoek was om thematische inzichten te verkrijgen over vestigingspatronen van innovatieve startups. Aangezien de aantallen observaties in buurten voor sommige topsectoren erg klein zijn, is voorzichtigheid geboden bij het trekken van conclusies. De sectoren Chemie en Tuinbouw zijn slechts in vijftien buurten vertegenwoordigd waardoor enkele uitschieters al snel tot vertekeningen zullen leiden. Plots van de thematische resultaten zijn te vinden in de figuren B.3.8 tot en met B.3.17 in bijlage B.3.3. De thematische verschillen per variabele worden eerst kort besproken, gevolgd door een meer gedetailleerde uitleg van het ruimtelijke patroon per thema.

Te beginnen met het aantal bestaande bedrijven, is opvallend dat de statistische waarschijnlijkheid van een LS&H-startup het meest toeneemt (+8%, $p < 0,001$) naarmate het aantal bedrijven in een buurt met een eenheid van tien toeneemt. Dit effect is naar verhouding het kleinste voor startups in de Tuinbouw (+4%, $p = 0,063$) en Logistiek (+5%, $p < 0,001$).

De omgevingsadressendichtheid houdt een positief significant verband met het aantal startups binnen de Creatieve Industrie en een negatief significant verband met het aantal startups in de Tuinbouw. Een toename van 1.000 adressen wordt geassocieerd met een 1,31 (+31%, $p = 0,032$) maal groter aantal startups per buurt in de Creatieve

Industrie terwijl de kans op het aantal startups in de Tuinbouw bijna zeven (-85%, $p = 0,076$) maal kleiner wordt.

Het aantal hoogopgeleiden in de gemeente is weinig significant. Alleen in de topsector Agri & Food zal het aantal startups fors lager uitvallen (-85%, $p = 0,053$) bij een toename van het aantal hoogopgeleiden in de gemeente van 100.000. Dit hoeft niet te betekenen dat het opleidingsniveau een rol speelt, maar wel dat er specifieke gemeenten zijn waar dit thema oververtegenwoordigd is.

Het aantal cafés is een significant verklarende variabelen binnen vier van de 10 thema's. Voor startups in de ICT (-19%, $p = 0,011$), HTSM (-22%, $p = 0,027$), Creatieve Industrie (-26%, $p < 0,001$) en Energie (-35%, $p = 0,001$) geldt dat er bij een hoger aantal cafés binnen een kilometer afstand een lager aantal startups wordt verwacht. Dit is vooral opvallend voor de Creatieve Industrie en ICT aangezien het in tegenstrijd is met de bevindingen van Florida & King (2016) die juist startups in de software en media sectoren associëren met binnensteden.

Goede bereikbaarheid over de weg was voor vijf van de tien thema's een significante voorspeller. Startups in de sectoren Agri & Food, HTSM, LS&H, en Logistiek waren vooral in de eerste en meest bereikbare categorie (afstand van 0 tot 1,5km tot oprit hoofdverkeersweg) te vinden aangezien er sprake was van een significante voorspelde afname (respectievelijk: -44%, $p = 0,033$; -50%, $p = 0,010$; -38%, $p = 0,056$; -46%, $p = 0,049$) bij de tweede (afstand van 1,5 tot 3km tot oprit hoofdverkeersweg) ten opzichte van de eerste categorie met de hoogste bereikbaarheid. De IRR-waarden van 0,00 bij de vierde ten opzichte van de derde categorie geeft aan dat er voor de meeste topsectoren geen startups gevestigd zijn in de meest afgelegen gebieden (afstand van meer dan 5,5km tot de hoofdverkeersweg). Voor de topsector Water geldt dat bereikbaarheid een significante verklarende rol speelt aangezien het model een afname (-82%, $p = 0,083$) voorspelt bij categorie 3 (3,5 tot 5,5km afstand tot hoofdverkeer) ten opzichte van categorie twee.

Bij de meeste startups is een belangrijk overstapstap niet verder dan 10 kilometer (gemeten over de weg) verwijderd. Er zijn onder de thema's geen significante verschillen gevonden tussen de eerste (0-10km) en tweede (10-20km) categorie. Voor startups in de Creatieve Industrie bleken vooral

de meest afgelegen categorieën drie en vier een significant negatief effect (respectievelijk -43% , $p=0,021$ en -61%, $p=0,058$) te hebben op het aantal startups. Voor startups in de sector Energie bleek een tegenovergesteld effect, met een significante toename van het aantal startups in categorie drie ten opzichte van categorie twee (+73%, $p=0,089$).

Zowel het aantal westerse allochtonen in de buurt als in de wijk is betrokken in de analyse. Voor de meeste topsectoren zijn significante resultaten gevonden. Opvallend is dat voor westerse allochtonen in de buurt significant negatieve effecten werden geconstateerd, terwijl dit voor westerse allochtonen in de wijk ging om positieve significantie. Verklaring hiervoor is dat op wijkniveau de correlatie met de OAD en het aantal bestaande bedrijven vrij sterk is waardoor deze positieve verbanden waarschijnlijk duiden op de aanwezigheid van urbanisatie-economieën. Op buurtniveau komt de heterogene verdeling van westerse allochtonen binnen de wijken aan het licht. Het negatieve verband lijkt daarom vooral te maken te hebben met het feit dat veel startups niet in woonbuurten maar bijvoorbeeld in bedrijventerreinen of binnensteden zijn gelokaliseerd waar relatief minder mensen wonen.

Op basis van de literatuur werd eerder al geconstateerd dat de groep westerse allochtonen sociaal economisch gezien mogelijk geen homogene groep zal zijn. Met name laag geschoolde Poolse arbeidsmigranten verschillen aanzienlijk van hoger geschoolde arbeidsmigranten afkomstig uit de rest van de westerse wereld. Helaas is dit onderscheid niet in de beschikbare data verwerkt. Opvallend is dat juist bij sectoren waar technisch geschoolde arbeidskrachten nodig kunnen zijn, de kans op een startup het meeste afneemt bij een toename van westerse allochtonen in de buurt (Agri & Food, -42%, $p<0,001$; LS&H -38%, $p<0,001$; Chemie -36%, $p=0,027$; HTSM -33%, $p<0,001$). Dit kan te maken met het feit dat startups in technische sectoren minder vaak in woonbuurten te vinden zijn.

Een duidelijker beeld, wordt verkregen door ook naar de onderlinge verschillen op wijkniveau te kijken. Opvallend is dat het model nu wel een relatief verhoogde voorspelling van het aantal HTSM (+222%, $p<0,001$) en LS&H (+171%, $p=0,004$) startups doet bij een toename van 1.000 westerse allochtonen in de wijk. Alleen de Tuinbouw⁴ overtreft deze score met een hogere IRR (+321%, $p=0,031$).

Dongen et al. (2014), die namens het PBL de relatie tussen vestigingsplaatsfactoren en de concentratie van topsectoren uitgebreid onderzoeken, hebben geen variabelen betrokken met betrekking tot migratieachtergronden of arbeidsmigranten. Zonder referentie kader is het lastig te zeggen of de gevonden effecten te maken zullen hebben met geschoolde of ongeschoolde arbeidsmigranten. Of misschien zelfs met beiden.

Aangezien startups in de Tuinbouw met weinig innovatieclaims (zie tabel 4.1) en een alternatief ruimtelijke patroon (zie figuur B.3.16) konden worden geassocieerd, is een mogelijke verklaring dat er sprake is van startups die bijvoorbeeld weinig innovatief maar wel arbeidsintensief zijn. Dit zou kunnen wijzen op laaggeschoolde westerse migranten die in de nabijheid van de startup wonen.

Tenslotte volgt de interpretatie met betrekking tot de invloedssferen van kennisinstellingen (of andere stedelijke aantrekkingskrachten) op nabijgelegen buurten binnen de bebouwde kom van de stad. Het gaat om een vergelijking tussen de onderlinge thema's en de twee steden Leeuwarden en Groningen. Te beginnen met buurten in de stad Groningen, springen twee thema's eruit: de kans op een Agri & Food startup zal volgens het model in de stad Groningen met bijna een factor tien toenemen (+967%, $p=0,008$) en voor de topsector Energie gaat het om een ruime vervijfvoudiging (+549%, $p=0,046$). De associaties met de overige thema's waren voor Groningse buurten niet significant. Stedelijke buurten in Leeuwarden hadden – in mindere mate – te maken met vergelijkbare thematische effecten voor startups in Agri & Food (+412%, $p=0,023$) en Energie (+407%, $p=0,012$). Voor de onderwijsinstellingen in beide steden kunnen opleidende faculteiten worden geïdentificeerd (Hogeschool Van Hall Larenstein [VHL], zd; Rijksuniversiteit Groningen [RUG], zd) die overeenkomen met beide thema's.

In Leeuwarden zijn ook voor andere thema's significante resultaten gevonden. De statistische kans op meer startups in de sectoren Chemie⁵ (+803%, $p=0,034$) en HTSM (+469%, $p=0,004$) waren relatief groter in Leeuwarden. Zowel Groningen als Leeuwarden beschikken over vergelijkbare opleidingen in de richtingen van deze thema's en de verschillen tussen de steden zijn dan ook niet heel

⁴ Dit kan vertekenen op basis van weinig waarnemingen.

⁵ Dit kan vertekenen op basis van weinig waarnemingen.

groot. Overige significante positieve associaties voor buurten in de stad Leeuwarden zijn gevonden voor LS&H (+399%, $p=0,010$), Logistiek⁶ (+357%, $p=0,048$), ICT (+214%, $p=0,093$) en Creatieve Industrie (+250%, $p=0,012$) maar verschillen niet heel veel met Groningen.

Na alle voorspellende variabelen te zijn langsge- lopen, wordt een algemeen beeld met vooral de opvallendheden van de vestigingspatronen per thema gesynthetiseerd. Aandacht gaat niet alleen uit naar onderlinge thematische verschillen maar ook naar afwijkingen ten opzichte van de theorie en het algemene topsector vestigingspatroon ge- schetst door Dongen et al. (2014).

Te beginnen met Agri & Food, in vergelijking tot de andere thema's is het meest opvallend dat de kans op startups relatief groter is in gemeente met minder hoogopgeleiden. Een mogelijke verklaring is dat een deel van deze startups juist gevestigd is nabij landelijke gebieden vanwege het nut of de noodzaak om activiteiten of innovaties toe te kun- nen passen in deze gebieden. Sommige van deze afgelegen gebieden zijn mogelijk uitsluitend ge- schikt voor economische activiteiten die agrarisch van aard zijn. Startups in de andere sectoren zullen op deze afgelegen plekken slechter vertegenwoor- digd zijn omdat ze er niets te zoeken hebben. Het patroon van de Agri & Food startups wordt ook ge- kenmerkt door de laagste incidentie ratio's in buur- ten en wijken waar westerse allochtonen wonen en een sterkere concentratie in de stedelijke gebied nabij de kennisinstellingen. Het deel van de Agro & Food startups dat juist in steden is gevestigd zal vanwege andere behoefte, zoals de aanwezig- heid van sociaal kapitaal, stedelijke voorzieningen of de bereikbaarheid (voor de werknemers zelf, zakenpartners of klanten), juist wel in steden zijn gevestigd. Als dit inderdaad biotechnologische startups betreft komt dit overeen met bevindingen van Florida & King (2016).

Voor de topsector Chemie zijn eigenlijk te weinig waarnemen om goed een ruimtelijk patroon vast te kunnen stellen. Opvallend is wel de concentratie rondom de kennisinstellingen die aanverwante op- leidingen in chemie (VHL, zd) en scheikunde (RUG, zd) aanbieden.

De Creatieve Industrie is de sector met de meeste

startups. In vergelijking met de andere topsecto- ren zijn startups in deze sector relatief het beste vertegenwoordigd in hogere stedelijke dichtheden. Toch bleek het aantal cafés – geassocieerd met stedelijke levendigheid – een significant negatief effect te hebben. Mogelijke verklaring hiervoor is dat startups in de Creatieve Industrie wel in stedelijke buurten maar niet in de binnensteden zijn gevestigd waar de huurprijzen hoger zijn en relatief minder plek is om te wonen. Startups in de Creatieve Industrie worden mogelijk makkelijker vanuit huis opgericht of verkieszen bedrijventer- reinen of andere vestigingslocaties buiten het hart van de stad. Dat Florida & King (2016) juist startups in de media sector associëren met binnensteden kan te maken hebben met een verschil tussen 'downtown Amerika' en de binnensteden van Noord-Nederland. Het kan ook te maken hebben met de brede definitie van de Creatieve Industrie in combinatie met het feit dat een groot deel van de Nederlandse mediabranche geconcentreerd is rond Amsterdam en Hilversum (Dongen et al., 2014), met als gevolg dat bedrijven binnen de Creatieve Industrie in Noord-Nederland naar ver- houding minder media-georiënteerd zijn en een ander locatiepatroon volgen.

Startups binnen de topsector Energie, waren rela- tief vaak te vinden in buurten ver verwijderd van belangrijke overstapstations met minder cafés binnen één kilometer. Mogelijk is een deel van de energie bedrijven gevestigd nabij plaatsen waar ook energie kan worden opgewekt, en dit is vaak niet in de buurt van cafés of overstapstations maar wel nabij het hoofdverkeer.

Voor startups in de topsectoren HTSM en ICT voorselt het poissonmodel per voorspeller een vrij gemiddelde verandering ten opzichten van star- tups in de andere sectoren. HTSM startups lijken niet in binnensteden (of nabij veel cafés) gevestigd te zijn maar eerder op goed bereikbare locaties nabij andere bedrijven rondom kennisinstellingen of andere stedelijke gebieden. ICT startups lijken geen specifieke voorkeuren te hebben ten opzichte van andere startups. Verklaring zou kunnen zijn dat de sector weinig gebonden is aan specifieke locatie-kenmerken.

LS&H-startups hebben de neiging zich dicht bij bestaande bedrijven en het hoofdverkeer te be- vinden. Afgaand op het negatieve verband met westerse allochtonen in de buurt, blijken LS&H startups relatief weinig in woonbuurten te zitten.

⁶ Voor Logistiek geldt t.o.v. Groningen geen grotere maar kleinere IRR.

Veel Logistieke startups lijken relatief weinig waarden te hechten aan hoogstedelijke of goed bereikbare locaties. Niet de meest rurale, maar wel meer suburbane buurten lijken de voorkeur te genieten.

Net als bij Chemie, geldt ook voor de Tuinbouw dat er weinig waarnemingen zijn. Toch wordt dit ruimtelijk patroon sterk gekenmerkt door afwijkingen ten opzichte van de andere thema's. Eerder zijn mogelijke aanwijzingen voor minder innovatieve activiteiten opgemerkt (zie tabel 4.1) en mogelijk ook arbeidsintensieve bedrijven gevonden. Dit zou kunnen verklaren waarom startups in de tuinbouw in buurten gevestigd zijn waar weinig andere bedrijven te vinden zijn.

Voor de topsector Water zijn op basis van de poissonmodellen geen opvallende significante verschillen ten opzichten van de andere thema's voorspelt.

4.2.1c Overige afwijkingen

Als laatste is gekeken of er nog opvallende afwijkingen konden worden geconstateerd op basis van cross-sectorale verbanden en de cases in validiteits-steekproef. In tabel 4.8 zijn de poissonmodellen met de X-sectorale variabelen en steekproef-controle resultaten als afhankelijke variabelen afgezet tegenover het totaal en de Top25 startups ter referentie.

Wat betreft de meest valide innovatieve startups (aangeduid als 'TOP') op basis van de steekproef-controle, was opvallend dat de OAD een negatief significant effect had in plaats van een positief effect. Dit komt overeen met studies die constateren dat economische verschillen tussen het stedelijke en rurale gebied op sommige plaatsen vervagen en dat traditionele manieren van innovatie meten minder goed in staat zijn deze innovatieve activitei-

ten op te merken (Renski, 2014; McCann & Ortega-Argilés, 2015; Naldi et al., 2015). Het lijkt er echter ook op dat de meer valide bevonden innovatieve startups ('TOP' innovatoren) vaker in woonwijken gevestigd zijn (in buurten met relatief veel westerse allochtonen). Dit kan te maken hebben met de vroege stadia van startupvorming, waarbij behaviourale locatie theorieën een grotere rol spelen en oprichters nog op (of nabij) hun woonadres gevestigd zijn. Ook viel het relatief grote belang van openbaar vervoer op voor de 'TOP' innovatoren uit de steekproef-controle.

Aangezien een aanzienlijk deel van de steekproef-controle voorkomt uit de geselecteerde cases met drie of meer cross-sectorale verbanden, verschillen de modellen met de cross-sectorale variabelen niet heel veel van de modellen met de cases uit de steekproef-controle. Bedrijven met minimaal 2 scores in het Top25 segment, waren relatief weinig te vinden in gemeenten met veel hoogopgeleiden maar juist wel relatief vaker in het stedelijk gebied van de onderwijsinstellingen in Leeuwarden en Groningen.

Als laatste opmerking moet worden benadrukt dat er over het algemeen sprake is van lage aantallen startups en kleine onderlinge verschillen die gevoelig zijn voor uitschieters en toevalligheden. Daar boven op wordt ook verwacht dat er veel ruis in de gegevens zit, omdat minder valide of ongeldige resultaten niet allemaal opgespoord en verwijderd konden worden. De resultaten schetsen wel een interessant beeld dat over het algemeen overeen komt met bevindingen uit de literatuur, of er kon een alternatieve verklaring gevonden worden. Het zou echter voorbarig zijn om definitieve conclusies te verbinden aan de ruimtelijke patronen die gevonden zijn.

Tabel 4.8 Quasi-poissonmodellen met cross-sectorale verbanden en de steekproef-controle groepen

	ALLES	Top 25	X-T25 som	X-T25 xT	X-T25 >1	X-T25 >1som		STEEK-PROEF	TOP-SEMI	TOP	SEMI	FLOP
<i>Voorspellers</i> (in buurt, tenzij anders vermeld)	IRR	IRR	IRR	IRR	IRR	IRR		IRR	IRR	IRR	IRR	IRR
(Intercept)	0,58 ***	0,15 ***	0,27 ***	0,42 ***	0,07 ***	0,17 ***		0,06 ***	0,04 ***	0,03 ***	0,02 ***	0,02 ***
Bedrijfsinschrijvingen 2016 (x10)	1,06 ***	1,06 ***	1,07 ***	1,06 ***	1,06 ***	1,06 ***		1,06 ***	1,07 ***	1,05 ***	1,08 ***	1,05 ***
Omgevingsadressendichtheid (x1.000)	1,14 **	1,12	1,14	1,07	0,75	0,80		0,97	0,90	0,66 •	1,05	1,08
Aantal hoogopgeleiden (in gemeente x100.000)	0,31 **	0,46	0,53	0,18	0,06 **	0,05 **		0,96	1,08	0,51	1,67	0,67
Aantal cafés binnen 1 km (x10)	0,75 ***	0,79 ***	0,76 ***	0,88	0,92	0,99		0,76 ***	0,76 *	1,00	0,66 *	0,79 **
Afstand tot oprit hoofdverkeersweg (dummy's in 4 categorieën)	1,2	0,72 ***	0,80 •	0,73 **	0,77	0,59 *	0,64 **	0,84	0,91	0,67	1,10	0,65
	2,3	0,65 **	0,80	0,58 **	0,38 **	0,30 **	0,72 **	0,78	0,96	0,94	0,97	0,38
	3,4	0,53 •	0,32	0,60 •	0,52	0,52	0,72	0,21 **	0,28 **	0,00	0,59	0,00
Afstand tot belangrijk overstapstation (dummy's in 4 categorieën)	1,2	0,76 **	0,82	0,93	0,76	1,25	1,38	0,84	0,88	0,36 **	1,41	0,00
	2,3	0,76 **	0,75	0,84	0,75	1,05	1,08	0,68	0,71	0,53 •	0,88	0,72
	3,4	0,65 **	0,54 *	0,60 •	0,52	0,52	0,45	0,21 **	0,28 **	0,00	0,59	0,60
Westerse allochtonen (x100)	0,77 ***	0,73 ***	0,74 ***	0,68 ***	0,75 ***	0,71 ***		0,76 ***	0,75 ***	0,83	0,71 ***	0,82 **
Westerse allochtonen (in wijk x1.000)	1,36 ***	1,35 *	1,27 **	1,30 •	1,81 ***	1,80 ***		1,63 *	1,67 *	2,28 ***	1,43 •	1,54 •
Groningen KI (dummy nabijheid)	0,1	3,10 ***	3,54 *	3,54 *	4,73 **	9,85 ***	9,89 *	2,42	2,18	1,57	2,69	3,10
Leeuwarden KI (dummy nabijheid)	0,1	1,99 ***	1,92 **	2,12 *	2,23 •	4,81 *	4,45 *	3,49 ***	3,53 *	6,84 ***	2,23	3,59 **
Observaties N / ≠ 0	1994 / 685	1994 / 285	1994 / 285	1994 / 194	1994 / 132	1994 / 132		1994 / 148	1994 / 111	1994 / 94	1994 / 72	1994 / 97
Cox & Snell's R2 / Nagelkerke's R2	0,61 / 0,69	0,19 / 0,33	0,41 / 0,50	0,38 / 0,41	0,12 / 0,29	0,31 / 0,41		0,13 / 0,31	0,09 / 0,27	0,04 / 0,21	0,07 / 0,27	0,04 / 0,23

• sig<0,1; *sig<0,05; **sig<0,01; ***sig<0,001

5 Conclusie & Discussie

In dit onderzoeksrapport is op basis van kwantitatief onderzoek en data-analyse getracht antwoord te krijgen op de vraag in hoeverre innovatieve startups met behulp van een *web-based* zoekmethode in beeld kunnen worden gebracht. Dit heeft een aantal methodische, geografische en thematische inzichten onthuld die na een samenvatting ter afsluiting van dit rapport worden geëvalueerd.

5.1 Conclusie

Deze studie heeft bedrijfswebsites gebruikt als bron voor innovatieonderzoek, en maakt deel uit van een onderzoekslijn die nog weinig in de economische geografie is toegepast. Specifiek ging de aandacht uit naar de 'online geografie van innovatieve startups' in Noord-Nederland. Daarmee komt deze studie tegemoet aan de uitdaging zoals geagendeerd door bibliografisch onderzoek over het ondernemende ecosysteem concept (Malecki, 2018) om potenties van verschillende innovatieve omgevingen beter aan het licht te brengen.

De studie was kwantitatief en exploratief van aard. Voordat de geografische resultaten worden besproken, wordt eerst ingegaan op de uitgevoerde methode van dataverzameling en de daarbij verkregen inzichten.

5.1.1 Identificatie innovatieve startups

Vergelijkbaar met het pionierswerk van Gök, et al. (2015) – en enkele voorlopers waaronder Youtie, et al. (2012), Arora, et al. (2013) en navolgers Kinne & Resch (2018) – zijn teksten van bedrijfswebsites onttrokken doormiddel van een geautomatiseerd *webscraping* proces. Er is gebruik gemaakt van data die door middel van de zoekmodellen uit de tool 'Innovatiespotter' van het bedrijf 'Q-Modus' zijn verkregen. Geïndexeerde bedrijfswebsites zijn geclassificeerd met relevantiescores op basis van tien innovatieve thema's. Deze thema's zijn weer gebaseerd op de Nederlandse topsectoren (Rijksoverheid, 2019). Tien datasets afkomstig uit tien deelthema's zijn gefilterd op startup kenmerken, gerangschikt op relevantiescores en vervolgens samengevoegd. Er is hoofdzakelijk naar de 25% hoogste relevantiescores gekeken omdat er

veel ruis aan minder relevante bedrijven werd verwacht bij lagere relevantiescores. Uit deze Top25 selectie zijn een aantal gerichte steekproeven getrokken om de kwaliteit van de automatische classificatie te beoordelen.

Op basis van de steekproef-controle werd slechts een kwart van de Top25 cases beoordeeld als zeer valide resultaat. Bij de helft bleek duidelijk geen sprake van een innovatieve startup te zijn. Dit valt echter vooral te wijten aan de operationalisering van startups als klein en jong bedrijf in combinatie met slecht onderhouden of verouderde registratiegegevens waar het filteringsproces op was gebaseerd. Ondanks het niet voldoen van bedrijven die na inspectie te groot of te oud zijn bevonden, zal een groot deel van deze afvallers toch dezelfde unieke kenmerken behouden waardoor de resultaten nog steeds als waardevol kunnen worden beschouwd voor de ruimtelijke analyse.

Bij zeventig procent van de gefilterde bedrijven bleek er op basis van websitegegevens daadwerkelijk sprake te zijn van topsector-expertise. Bij het grootste deel aan zogenaamde *false positives* bleken trefwoorden op de website in een andere context voor te komen. Deze vorm van ruis zou normaal gesproken bij lagere relevantiescores verwacht worden, maar kwam tegen de verwachting in ook voor bij de hoogste relevantiescores. Een logische verklaring hiervoor is dat de ruis bij hoge relevantiescores vooral werd veroorzaakt door misleidende webportalen – zoals nieuws- en vacaturesites die doorverwijzen naar topsector-informatie – en bedrijven met vooral veel klanten in de topsectoren zonder zelf over typische topsec-

tor-kennis te beschikken.

De meerderheid van de resultaten bleek echter niet alleen een goede voorspelling van bedrijven met topsector-expertise, maar ook een goede voorspeller voor (onderbouwde) innovatie claims op de bedrijfswebsites. Toegegeven, dit laatste biedt weinig garantie voor daadwerkelijke innovatieve output. Zoals bevestigd door de literatuur (Garcia & Calantone, 2002; Smith, 2005) is dit deels inherent aan innovatie aangezien het zich als multidimensionaal fenomeen (Godin, 2002) lastig laat kwantificeren. Rekening houdend met een bepaalde selectie, omdat bedrijven variëren in hun motivatie voor het plaatsen (of niet plaatsen) van informatie op websites, kan worden geconcludeerd dat de methode zoals uitgevoerd in deze studie – ondanks verbeterpunten die later worden besproken – interessante nieuwe inzichten biedt die niet of minder makkelijk worden verkregen met behulp van traditionele methoden.

De toegevoegde waarde wordt duidelijk door de aard van de resultaten te vergelijken met dat van gangbare methoden. Kwantitatieve onderzoeken gericht op output – bijvoorbeeld in de vorm van verkregen durfkapitaal, patenten of gerealiseerde groei – zijn niet in staat het oorspronkelijke groeipotentieel en het ongelijke belang van actoren en processen te monitoren of meten (Roundy, et al., 2017; Fazio, et al., 2016). Voor een volledig beeld van innovatie mag de output niet los worden gezien van de input. Bij kwantitatieve metingen van de input worden doorgaans R&D-indicatoren gebruikt, met als nadelig gevolg dat kleinere of minder technologisch georiënteerde bedrijven onterecht ondervertegenwoordigd zijn.

De *webscraping* methode gebruikt voor deze studie, heeft zich primair op de input gericht en is juist wél gericht op kleine bedrijven en niet uitsluitend op technologische innovatie. Het is echter wel een momentopname. Op zoek naar de innovatieve startup kan in deze studie daarom beter worden gesproken over het innovatiepotentieel aan startups omdat identificatie van innovatieve thema's geen garanties biedt voor innovatieve ontwikkelingen in de toekomst. De vergaarde gegevens – inclusief veel niet-technologische claims van innovatieve activiteiten door relatief kleine bedrijven – onderscheiden zich van het gebruikelijke beeld van innovatie. Het biedt daarmee nieuwe mogelijkheden voor benaderingen van innovatie- en groeipotentieel.

De groep gevonden innovatieve bedrijven zijn enerzijds homogeen omdat vergelijkbare aanwijzingen voor topsector-expertise en innovatie op hun websites te vinden zijn. Homogeniteit blijkt ook uit de vele cross-sectorale verbanden die binnen individuele bedrijven zijn geïdentificeerd. Op basis van de steekproefcontrole lijken deze interne crossovers ook verband te houden met een verhoogde aanwezigheid van innovatieve activiteiten. Anderzijds wijst de steekproefcontrole – ondanks de zojuist besproken gelijkenissen – op een sterk heterogene groep: Zowel topsectoren als innovatieve activiteiten zijn breed geformuleerde concepten, daarbinnen bestaan vele variaties. Binnen deze variatie is een belangrijk onderscheid de mate van onafhankelijkheid aangezien dit implicaties bevat voor de unieke (disruptieve) rol die innovatieve startups kunnen vervullen in de regionale economische dynamiek. Tegelijkertijd kan onafhankelijkheid ook een kwetsbaarheid zijn, een teken van behoefte aan meer ondersteuning.

Op basis van de meest valide steekproef resultaten bleek het merendeel van de innovatieve startups een *spin-off* uit het bedrijfsleven. Daarvan konden in enkele gevallen nauwe banden met een moederonderneming worden geïdentificeerd en bij een iets grotere groep waren aanwijzingen te vinden voor 'diversificatie'. Gezien vanuit evolutionaire theorieën zouden diversifiërende startups tot interessante innovatie kunnen leiden wanneer kennis vanuit ongelijksoortige industrieën succesvol wordt geïmplementeerd.¹ Er waren ook enkele innovatieve *spin-offs* uit kennisinstellingen waarvan de implicaties vergelijkbaar zijn met *spin-offs* uit moederbedrijven. Het resterende deel aan innovatieve startups leek op basis van online informatie zoals de LinkedIn-profielen van oprichters naar verhouding het meest onafhankelijk te zijn. Deze groep bestond uit veel jonge starters (veel alumni) maar ook meer 'doorleefde' oprichters die hun ideeën konden baseren op levenservaring of -overtuiging die niet in verband gebracht kon worden met betekenisvolle werkervaringen uit het verleden.

Ter besluit van het eerste deel van de conclusie, en in lijn met de redenering van Baumol (2010), zal

¹ Gerelateerde diversiteit zou leiden tot meer innovatie (Mowery et al., 1996; Boschma & Lammarino, 2009; Boschma et al., 2014; Janssen, 2015) en ongerelateerde diversiteit zou nog belangrijker kunnen zijn en tot disruptieve innovatie kunnen leiden (Nooteboom et al., 2007; Enkel & Gassmann, 2010; Castaldi et al., 2014).

het innovatiepotentieel aan startups dat gevonden is in het onderzoeksgebied, wat betreft afzonderlijke strategieën niet homogeen zijn. Zo zijn verschillende startups in de steekproef-controle als niet valide bestempeld wanneer deze alleen bestaande ideeën lijken te kopiëren. Daar bovenop zullen innovatieve startups die wel claimen zelf vernieuwend te zijn zich mogelijk ook schuldig maken aan 'parasitair' gedrag en geen wezenlijke bijdrage leveren aan innovatie.² Alleen een klein deel zal betekenisvolle innovaties introduceren en slechts een paar innovatieve startups zullen baanbrekend zijn.

5.1.2 Ruimtelijke en thematische inzichten

Deze studie presenteert ook een voorspellingsmodel voor de locaties van de zojuist besproken bedrijfsobservaties waarvan geconcludeerd is dat deze in meer en mindere mate gekarakteriseerd kunnen worden als innovatieve startups. Voor de modellen zijn poissonregressies toegepast en is gebruik gemaakt van data afkomstig van het CBS.

Locatiepatronen zijn geïdentificeerd door eerst verkennend onderzoek op buurtniveau uit te voeren naar startups die in verband zijn gebracht met verschillende innovatieve thema's of sectoren in een zogenaamde *exploratory spatial data analysis* (ESDA). Op basis van deze verkenning, de literatuur en beschikbaarheid van data zijn verklarende variabelen samengesteld.

Het voorspellingsmodel leek goed te werken, maar kende ook beperkingen (zie discussie 5.2.1b). De resultaten kunnen worden toegeschreven aan verschillende geografische concepten afkomstig uit de literatuur over het ondernemende ecosysteem (Malecki, 2018), en daarmee samenhangende geografische theorieën zoals locatietheorieën (Atzema, et al 2002).

Urbanisatie-economieën (neoklassieke locatietheorie) zijn veel voorkomende en bestudeerde determinanten als het gaat om locatiepatronen van (innovatieve) bedrijven (Jacobs, 1961; Marshall, 1890; Glaeser et al., 1992; Arauzo-Carod et al., 2010). Om hier in het model voor te controleren is de omgevingsadresdichtheid (OAD) in het model betrokken als indicator voor stedelijke dichtheid en diversiteit. Omdat urbanisatie-economieën veelomvattend zijn, waren verwante deelaspecten

die correleren met de OAD ook opgenomen in de analyse. Het gaat om het aantal bedrijven, cafés, westerse allochtonen in de buurt en wijk, de nabijheid van vervoersinfrastructuren, kennisinstellingen en hoogopgeleiden in de gemeente. Omdat de analyse hoofdzakelijk is uitgevoerd op het micro-schaalniveau van de buurt, waren gegevens zoals het aantal bedrijven minder gecorreleerd met de OAD dan bij hogere aggregaties het geval zou zijn.³ Het betrekken van aspecten die correleren met de OAD heeft er voor gezorgd dat de OAD lager uitvalt en een soort controle variabele is geworden voor 'overige stedelijke dichtheid of diversiteit'.

Binnen de automatisch geclassificeerde resultaten, is aan de hand van de voorspellingsmodellen een geografische vergelijking getrokken tussen de hoogst scorende resultaten en de gehele groep resultaten. Ondanks het feit dat de laagste relevantiescores niet uitgebreid zijn onderzocht, kon worden aangenomen⁴ dat aanwijzingen voor innovatieve activiteiten op bedrijfswebsites sterk afnemen bij lagere relevantiescores. Op basis van deze redenering zijn de hoogst scorende resultaten geassocieerd met innovatievere activiteiten in vergelijking tot de gehele groep. Enkele verschillen zijn naar voren gekomen waarvan de meest interessante suggereert dat innovatievere startups eerder in meer levendige gebieden (meer cafés in de omgeving) te vinden zijn. Dit is vergelijkbaar met empirische bevindingen van Malizia & Motoyama (2018) en de opvattingen van Jacobs (1961, 1970) en Florida (2002) over creatieve en levendige stedelijke locaties die innovatieve en startup-formatie zouden stimuleren. Daarnaast bleek op basis van de vergelijking dat innovatievere startups Groningen mogelijk verkiezen boven Leeuwarden. Dit zou mogelijk te maken kunnen hebben met de rol van Groningen als universiteitsstad (Guerrero et al., 2016) maar ook met superieure urbanisatie-economieën waaronder netwerk-mogelijkheden (Bathelt et al., 2004; Gauthier et al., 2018; Malecki, 2018).

Op basis van de steekproef-controle en in vergelijking tot andere sectoren was opvallend dat

³ Zie Kinne & Resch (2017) die laten zien dat de correlatie tussen bedrijfsaantallen en populatieaantallen tussen het aggregatieniveau van 1 en 5 kilometer verdubbelt.

⁴ Op basis van simpele checks, uitgebreide ervaringen van de Innovatiespotter (A. Denekamp, persoonlijke communicatie, 9 september 2018), methodologische logica en significante controleresultaten van uitgevoerde ANOVA-tests binnen de hoogst scorende resultaten.

² Opgemerkt moet worden dat, voor een succesvolle introductie van innovatie, zogenaamde *early adopters* ook een belangrijke of zelfs essentiële rol kunnen spelen.

op bedrijfswebsites van startups in de Creatieve Industrie, relatief de meeste innovatieclaims te vinden waren. Het volledige automatische geclassificeerde innovatiepotentieel aan startups in de Creatieve Industrie bleek daarnaast als enige sector een significant positief verband met de OAD te houden. De overige sectoren werden gekenmerkt door een positief maar niet significant verband. De verbanden komen overeen met theorieën die stedelijke diversiteit beschouwen als aanjager van innovatie en startup-formatie (o.a. Jacobs, 1961; Desrochers & Sautet, 2008; Glaeser 2011).

Zeer opvallend was dat de meest valide resultaten op basis van de steekproef-controle een duidelijk tegenovergesteld significant verband tussen het aantal innovatieve startups en de OAD lieten zien. Een mogelijke verklaring hiervoor is dat het automatisch geclassificeerde innovatiepotentieel ruis bevat van ondernemingen die al in een later stadium van groei verkeren of meer traditionele activiteiten ontplooiën. Hierdoor spelen (neo)klassieke locatietheorieën (Atzema et al., 2002) mogelijk een relatief grotere rol. Deze theorieën hebben betrekking op specifieke locatiekenmerken die productieprocessen optimaliseren (klassieke locatietheorie) of concurrentieposities verstevigen (neoklassieke locatietheorie).

Bij de valide bevonden resultaten uit de steekproef-controle is met grote zekerheid vastgesteld dat het om kleine en jonge startups gaat. Deze startups bevinden zich in de vroege stadia van startupvorming. Op basis van theoretische verwachtingen (Holden, & Pen, 2004; Arauzo & Manjon, 2004; Stam, 2007). werd voorspeld dat behaviourale locatietheorieën in deze vroege fase een grotere rol zouden spelen. Deze verwachting lijkt te zijn uitgekomen: ondernemers in de vroege stadia van startup-formatie lijken eerder geneigd dicht bij hun cognitieve of fysieke oorspronkelijke locatie (bijvoorbeeld woonadres) te blijven. Significante voorspellingen van het locatiekeuze-model associëren deze startups met buurten waarin relatief veel westerse allochtonen wonen. Dit zijn vermoedelijk woonbuurten met een lagere OAD dan bijvoorbeeld de binnenstad.

De bovengenoemde resultaten die aangeven dat valide innovatieve startups zich in relatief lagere stedelijke dichtheden bevinden, sluiten ook aan bij andere stromingen in de literatuur die aantonen dat economische verschillen tussen stedelijke en landelijke gebieden op sommige plaatsen vervagen en dat traditionele manieren van innovatie-metingen

(voornamelijk gericht op technologische innovaties gebaseerd op patenten en R&D-indicatoren) minder goed in staat zijn om dit op te merken (Renski, 2014; McCann & Ortega-Argilés, 2015; Naldi et al., 2015). Het kan ook te maken hebben met het feit dat goedkope werkplekken aan de stadsranden op bedrijventerreinen in lage dichtheden aantrekkelijker zijn in de opstartperiode van een startup.

Naast voorgaande bevindingen zijn er voor de afzonderlijk bestudeerde topsectoren verschillende ruimtelijke patronen geïdentificeerd. Alleen de meest opvallende en afwijkende resultaten worden hieronder samengevat.

Ten eerste waren er aanwijzingen – en deze werden versterkt door bevindingen op basis van een handmatig uitgevoerde steekproefcontrole – dat sommige topsectoren beter hadden kunnen worden opgedeeld in categorieën op basis van hoofd- of deelactiviteiten die gepaard zouden gaan met verschillende locatievoorkeuren. Een dergelijk onderscheid zou enkele van de gevonden verklaringen kunnen versterken. Ook zou het meer inzicht kunnen bieden om theorieën over gerelateerde diversiteit (zie voetnoot 1) te ondersteunen.

Het algehele locatie patroon van startups in de Agri & Food sector bleek bijvoorbeeld gekenmerkt door een voorkomen in zowel landelijke als stedelijke gebieden inclusief concentraties rondom kennisinstellingen. Een aannemelijke verklaring zou zijn dat een deel van deze startups juist gevestigd zijn nabij landelijke gebieden om er ter plekke activiteiten of innovaties toe te kunnen passen terwijl een ander deel juist baat heeft bij stedelijke voorzieningen zoals laboratoria of betere toegang tot sociaal kapitaal. Dit laatste strookt met verwachtingen op basis van de literatuur (Alvarez-Garrido & Dushnitsky, 2016; Florida & King, 2016; Colombo et al., 2017). De poissonmodellen bevestigen in ieder geval de verklaring voor het aandeel Agri & Food startups in de stad, aangezien het locatiekeuzemodel significante voorspellingen deed voor een relatief hoog aantal Agri & Food startups nabij kennisinstellingen en een relatief laag aantal in woonwijken (buurten waar relatief veel westerse allochtonen wonen). Specifieke locatiekenmerken in het landelijk gebied konden door de modellen niet worden herkend.

Niet alleen waren er op basis van de steekproef-controle het meeste innovatieclaims opgemerkt bij startups in de Creatieve Industrie, het is ook de sector met het grootste aantal bedrijven in ver-

gelijking tot andere sectoren waardoor de onderzoeksresultaten statistisch gezien betrouwbaarder zijn. Daarnaast is deze sector interessant omdat innovatieonderzoek vaak gericht is op sterk technologische sectoren (Godin, 2002) en onderzoek naar innovatie in kleine creatieve bedrijven schaars is (Camelo-Ordaz et al., 2012). Ondanks het gevonden sterke verband tussen het innovatiepotentieel aan startups in de Creatieve Industrie en de OAD van de buurt, bleek het aantal cafés een significant negatief effect te hebben. Een verklaring ondersteund door de literatuur is dat het effect van de OAD niet-lineair is maar een omgekeerd U-vormig profiel volgt (Arauzo-Carod et al., 2010). Dit betekent dat vanaf een bepaalde drempel negatieve economische effecten optreden, zoals een te hoge huurprijs in binnensteden waar wel cafés te vinden zijn maar minder woningen (die mogelijk als eerste registratieadres van een startup fungeren) en ook minder betaalbare bedrijfspanden te vinden zijn. Het beeld van Jacobs (1961, 1970) en Florida (2002) over creatieve en levendige stedelijke locaties als bron voor startup-formatie hoeft dus niet te worden ontkracht, in tegendeel, het zou goed waar kunnen zijn.

De gevonden resultaten en verklaringen van de resterende ruimtelijke patronen per topsector zijn vergelijkbaar met verklaringen die al gegeven zijn. Zo is het ruimtelijke patroon van Energie vergelijkbaar met Agri & Food aangezien opvallende resultaten in meer rurale gebieden suggereren dat er behoefte is aan 'landelijke' of goedkopere ruimte voor bedrijfsactiviteiten terwijl meer stedelijke resultaten suggereren dat urbanisatie-economieën een rol spelen. Rurale afwijkingen in het patroon zijn ook gevonden voor de Tuinbouw, maar in te-

genstelling tot de Agri & Food sector zijn er weinig aanwijzingen voor innovatieve activiteiten bij deze bedrijven.

Voor het overige innovatiepotentieel binnen de sectoren HTSM, ICT, LS&H, Logistiek en Chemie lijken vooral urbanisatie-economieën een rol te spelen. Daarbij zijn HTSM, Chemie en in mindere mate LS&H startups sterker geconcentreerd rondom kennisinstellingen op goed bereikbare locaties, bij bereikbaarheid geldt voor LS&H juist een relatief groter belang. ICT startups lijken op basis van de voorspellers in het model geen sterk onderscheidende voorkeuren te hebben en komen deels overeen met het ruimtelijke patroon van de Creatieve Industrie. Verklaring zou kunnen zijn dat de sector weinig gebonden is aan specifieke locatie-kenmerken en ook dat de sector veel overlapt met de Creatieve Industrie. Ten slotte lijkt het innovatiepotentieel aan startups in de Logistiek relatief vaker in suburbane buurten te zitten. Mogelijk is voor deze suburbane bedrijven vooral goedkope bedrijfsruimte van belang.

Dit tweede concluderende gedeelte eindigt met de constatering dat de geografische analyse op buurt-niveau met unieke gegevens over startup-vorming en innovatieve activiteiten die normaliter onder de radar blijven – ondanks ruis in de data – nieuwe inzichten heeft onthuld. Voor een groot deel worden theoretische en empirische verwachting door de data bevestigd, in sommige gevallen was er op voorhand weinig empirisch vergelijkingsmateriaal. Bestaande vermoedens en theorieën over creatieve stedelijke locaties zijn consistent met de resultaten uit deze studie, maar meer en uitgebreider onderzoek zou dit moeten bevestigen.

te vullen. De toepassing van de methode in deze studie was bij sommige aspecten nog wel problematisch en laat genoeg ruimte voor verbetering.

Een aanzienlijk gebrek was dat zoekmodellen in dit onderzoek niet zijn samengesteld door de onderzoeker maar als secundaire ruwe data zijn verkregen en verwerkt. Dit lijkt efficiënt, maar het niet zelf specificeren en ontwikkelen van de zoekmodellen vormde een beperking. Resultaten van zoekmodellen konden pas laat in het proces worden verkregen, bovendien konden ze niet eenvoudig worden aangepast om specifiek startups te filteren aangezien gekoppelde registratiegegevens veel

5.2 Discussie

In deze sectie wordt dieper ingegaan op de bruikbaarheid van de *webscraping*-methode en de validiteit en betrouwbaarheid van de ruimtelijke resultaten. Daarnaast wordt er gereflecteerd op het onderzoeksproces. Ook worden kansen voor toekomstig onderzoek benadrukt.

5.2.1 Sterke en zwakke punten

5.2.1a Webscraping en automatische classificatie

Zoals verwacht is *webscraping* als opkomende methode meer dan geschikt bevonden om enkele bestaande lacunes in geografisch innovatieonderzoek

fouten bevatten. Gebrek aan afstelling resulteerde niet alleen in onterecht gevonden resultaten (*false positives*), maar ongetwijfeld ook in onterecht niet gevonden innovatieve startups (*false negatives*) die nog tussen de ruis konden zitten of niet door het filteringsproces zijn gekomen. Voor deze problemen kunnen relatief eenvoudige en uitgebreidere oplossingen worden aangedragen.

Een relatief simpele oplossing zou kunnen zijn, om toegang te krijgen tot nauwkeurigere gegevens van bedrijven. Het is bekend dat aanvullende gegevens van het CBS of het LISA-arbeidsregister alternatieve methoden hanteren om gegevens te corrigeren of bij te schatten. Deze gegevens zijn momenteel onder strenge voorwaarden en of tegen betaling beschikbaar, maar bieden nog steeds geen garantie op registratiegegevens zonder onnauwkeurigheden. Wanneer gegevens als open data beschikbaar worden gesteld, kleven daar ook nadelen aan. Ongewenste commerciële benaderingen moeten zoveel mogelijk worden vermeden. Daarnaast zorgt de politieke vraag of de belastingbetaler of de gebruiker moet betalen voor het gebruik van bedrijfsgegevens voor aanvullende controversie (KvK, 2019).

Een alternatief voor gebrekkige handelsregisters, is het aanpassen van de zoekmodellen zodat deze naast herkenning van topsectoren ook gericht kunnen worden op herkenning van bedrijfs- startup- of andere kenmerken die de validiteit bevorderen. Uit dit onderzoek is gebleken dat niet alleen websites van bedrijven, maar ook extra online bronnen zoals LinkedIn en de Wayback Machine (Kahle, 2018) hiervoor geschikt kunnen zijn.

Een uitgebreidere oplossing voor het validiteitsprobleem, is om grotere veranderingen door te voeren in de zoekmodellen. Wederom kan hier een keuze worden gemaakt tussen eenvoud en complexiteit. Het gevaar van een sensitievere zoekopdracht is dat uitsluiting van *false positives* gepaard kan gaan met een toename in *false negatives*. Er bestaat echter veel keuzevrijheid om zoekmodellen nader te specificeren en combineren waardoor relevantiescores meer in overeenstemming kunnen worden gebracht met de beoogde resultaten. *False positives* waren nu nog te veel te vinden tussen de hoogste scores.

Rigoreuzere aanpassingen van zoekmodellen vereisen veel voorbereidend werk. Verschillende studies hebben echter aangetoond dat gedeeltelijk

gecontroleerde en geautomatiseerde leerprocessen de nauwkeurigheid van datasets aanzienlijk kunnen verhogen (o.a. Gök, et al., 2015; Roelands, van Delden, & Windmeijer, 2017; Kinne & Resch, 2018). Investeren in geavanceerdere zoektechnologie gebaseerd op ontwikkelingen in het domein van *machine learning* – vooral met behulp van zogenaamde neurale netwerken die op grote gegevensbronnen werken (Nilsson, 1998, 2015) – kan de moeite waard zijn. Dit laatste geldt bij uitstek voor projecten die optimaal gebruik maken van de schaalbaarheid en herhaalbaarheid van geautomatiseerde processen. Voor zowel onderzoek naar startups als innovatie schuilt hierin een grote kans om een leemte in innovatieonderzoek binnen de economische geografie op te vullen. Er blijkt een grote behoefte aan longitudinaal onderzoek dat inzicht biedt in zowel de input als de output rondom innovatie.

Naast de schaalvoordelen voor uitbreiding van onderzoeksresultaten, kan de kracht van *webscraping* worden samengevat als: het versterken van bestaande gegevens; in een korter tijdsbestek; opgesplitst in meer categorieën; tegen lagere kosten; en met minder respondentenbederf (zie ook Hackl, 2016; Hessisches Statistisches Landesamt, 2018).

Toch is het opvallend dat *webscraping* bij lange na niet wijdverbreid is binnen de wetenschappelijke literatuur van de economische geografie. Wel zijn er signalen dat projecten in toenemende mate van de grond komen. Mogelijk heeft de trage opname van *webscraping* te maken met het angstige sentiment rondom big data, privacy en eventueel misbruik ervan. Ook heerst er waarschijnlijk scepsis over de betrouwbaarheid van innovatie onderzoek dat alleen berust is op websites van de bedrijven zelf. Dit is onterecht omdat *webscraping* altijd als aanvullende en niet als een vervangende methode moet worden gezien. De meest overtuigende verklaring is echter dat technische knowhow, hulpmiddelen en de wil om in deze alternatieve technieken te investeren ontbreekt. Het bestaan van een dergelijke barrière zou ook verklaren waarom vooral statistische bureaus (CBS, 2018; Hessisches Statistisches Landesamt, 2018) met meer middelen of ondoorzichtige⁵ bedrijven uit de private sector (Orb Intelligence, 2018; CrunchBase, 2018; Dealroom, 2018) zich hier wel mee bezig houden om het gat

⁵ Met ondoorzichtig wordt bedoeld dat de bedrijven hun methoden niet openbaar hebben gemaakt waardoor de werkwijze onduidelijk blijft voor buitenstaanders. Hier is verder geen onderzoek naar gedaan.

op te vullen. Indexrapporten zoals die van Startup Genome (Gauthier et al., 2018) baseren zich hierop en vervolgens haken wetenschappelijke studies (Malecki, 2018) hier op aan.

Toch is het begrijpelijk dat er wordt afgezien van deze nieuwe technieken omdat succesvolle toepassing van automatische processen zoals *webscraping* en automatische classificatie (met of zonder implementatie van *machine learning*) niet eenvoudig is en om een lange adem vraagt. Het niet gebruiken van *webscraping* als onderzoekstechniek versterkt ook de onbekendheid, omdat de mogelijkheid voor vele anderen onzichtbaar blijft.

5.2.1b Ruimtelijke technieken en resultaten

Net als bij veel andere geografische tellingen van bedrijven zijn poisson-regressies toegepast om het aantal startups per buurt als afhankelijke variabele te modelleren. Poisson-regressies worden in de economische geografie veel gebruikt in zogenaamde *count data models* (Arauzo-Carod et al., 2010). Ondanks interessante onthullingen van de modellen in deze studie, zijn enkele beperkingen gevonden waarvan de oplossingen het model verder zouden kunnen verbeteren.

Omdat automatische classificatie het mogelijk maakt op een kleinere schaal te meten, moet de benadering van modellering hier ook op aangepast worden. De correcte operationalisering van locatiefactoren werd echter moeilijker op het kleine schaalniveau van de buurt, omdat verschillende locatiefactoren andere effecten hebben op verschillende geografische schalen. Kinne & Resch (2017) halen als voorbeeld aan dat een levendig nachtleven een positief effect kan hebben op de aantrekkelijkheid van de stad, terwijl bedrijven nog steeds de voorkeur geven aan rustige buurten (resultierend in een negatieve invloed op een meer gedetailleerd schaalniveau). Voor het probleem van effecten op verschillende schaalniveaus bestaan geavanceerde statistische oplossingen (zie Westerholt, Resch, & Zipf, 2015; Arauzo-Carod & Manjón-Antolín, 2012) waarvan verwacht wordt dat deze betere resultaten opleveren dan de poisson-regressies die in deze studie zijn uitgevoerd.

Andere onopgeloste problemen hebben betrekking op mogelijke verstoringen als gevolg van ruimtelijke autocorrelatie (Anselin, 1988; Dormann et al., 2007; Lambert et al., 2010; Bivand & Wong, 2018) en overtollige nullen op gedetailleerde geografische niveaus (Liviano & Arauzo-Carod, 2013).

Deze studie was – ondanks pogingen die gedocumenteerd zijn in bijlage B4 – echter niet ontworpen om geavanceerde statistische modellen tot een oplossing te brengen. Dit betekent dat het ruimtelijke regressiemodel relatief eenvoudig is gebleven.

Naast ontoereikende statistische technieken, hebben de regressiemodellen ook ingeboet aan voorspellingskracht door een gebrek aan beschikbare variabelen of bewerking van bestaande gegevens tot meer bruikbare indicatoren. Deze problemen worden wederom versterkt doordat sommige gegevens alleen beschikbaar zijn op hogere schaalniveaus. Locatiequotiënten van bedrijven konden alleen op basis van het schaalniveau van de gemeente worden berekend. Pogingen dit effect te herleiden naar het buurtniveau waren niet succesvol.

Ten slotte heeft deze studie uitgebreid de literatuur besproken op basis van het concept van een ondernemend ecosysteem. Dit concept is veelomvattend maar verhoudingsgewijs minder gericht op fysieke determinanten en meer op de kracht van netwerken en individuele actoren in het ondernemende systeem. Het belang van netwerken en cruciale actoren zoals belangrijke sleutelfiguren of organisaties moet worden onderkend. Buczkowska & De Lapparent (2014) erkennen ook het belang van ondernemersoriëntatie binnen de bevolking, en dat er meer werk nodig is om genuanceerde methoden voor kwantitatief onderzoek verder te ontwikkelen. Szerb et al. (2018) laten zien dat kwantitatieve effecten wel degelijk gemodelleerd kunnen worden. Dit is echter niet gemakkelijk en vereist toereikende middelen en toewijding om variabelen te betrekken en te beheren over complexere sociaal-culturele en financiële concepten.

Deze discussie eindigt met de notie dat nieuwe benaderingen om alternatieve krachten van innovatie te detecteren het huidige debat kunnen verscherpen. Lucas & Fuller (2017) pleiten zelfs voor een nieuwe kijk op regels binnen de institutionele context om meer innovaties met maatschappelijke waarde mogelijk te maken. Onafhankelijke startups met innovatieve ideeën lijden aan een structureel gebrek aan ondersteuning (Wymer & Regan, 2005; Faems & Wennink, 2018). Door een beter beeld van deze groep startups en ideeën te genereren kunnen publieke leiders hier beter op anticiperen. Voorbeelden van ondersteuning die publieke leiders beschikbaar kunnen stellen, zijn: toegang tot specifieke kennis of tot kapitaal en het wijzen op relevante connecties, bronnen van talent, potentiële

risico's of juridische kwesties. Daarnaast zouden longitudinale studies over het innovatiepotentieel niet alleen beter richting kunnen geven aan beleid maar ook langdurig monitoring mogelijk maken waardoor beleidsevaluaties meer opleveren. Dit zou aan bij kunnen dragen dat het structurele gebrek aan tastbare en ontastbare hulpbronnen niet langer de ontwikkeling van veel goede ideeën belemmert. Goede ideeën, zo werd al in de inleiding opgemerkt, zijn hard nodig in een moderne economie (zie Europese Commissie, 2013; Pan & Yang, 2018). Aangepast beleid zou niet alleen beter kunnen anticiperen op goede ideeën, het zou ook minder goede ideeën kunnen terugdringen ten gunste van nog waardevollere nieuwe ideeën.

Referenties

- Acs, Z. (2006). How is entrepreneurship good for economic growth? . *Innovations*, 1, 97– 107. doi: 10.1162/itgg.2006.1.1.97.
- Acs, Z., Stam, E., Audretsch, D., & O'Connor, A. (2017). The lineages of the entrepreneurial ecosystem approach. *Small Business Economics*, 49(1), 1-10.
- Agarwal, R., Echambadi, R., Franco, A., & Sarka, M. (2004). Knowledge transfer through inheritance: Spin-out generation, development, and survival. *Academy of Management journal*, 47(4), 501-522.
- Alvarez-Garrido, E., & Dushnitsky, G. (2016). Are entrepreneurial venture's innovation rates sensitive to investor complementary assets? Comparing biotech ventures backed by corporate and independent VCs. *Strategic Management Journal*, 37(5), 819-834.
- Annoni, P., Dijkstra, L., & Gargano, N. (2017). The EU Regional Competitiveness Index 2016. Working Papers, European Union.
- Anselin, L. (1988). *Spatial Econometrics: Methods and Models*. Dordrecht: Kluwer Academic Publishers.
- Arauzo, J., & Manjon, M. (2004). Firm Size and Geographical Aggregation: An Empirical Appraisal in Industrial Location. *Small Business Economics*, 22, 299–312.
- Arauzo-Carod, J., Liviano-Solis, D., & Manjón-Antolín, M. (2010). Empirical studies in industrial location: an assessment of their methods and results. *Journal of Regional Science*, 50(3), 685-711.
- Arauzo-Carod, J., & Manjón-Antolín, M. (2012). (Optimal) spatial aggregation in the determinants of industrial location. *Small Business Economics*, 39(3), 645-658.
- Ardilley, P., Audric, S., Bellefon, M., Buron, M., Durieux, E., Eusebio, P., . . . Sémécurbe, F. (2018). *Manuel d'analyse spatiale*. Insee Eurostat 2018.
- Arora, S., Youtie, J., Shapira, P., Gao, L., & Ma, T. (2013). Entry strategies in an emerging technology: A pilot web-based study of graphene firms. *Scientometrics*, 95(3), 1189–1207. doi:10.1007/ s11192-013-0950-7.
- Arthur, W. (1994). *Increasing returns and path dependence in the economy*. University of michigan Press.
- Arzaghi, M., & Henderson, J. (2008). Networking off madison avenue. *The Review of Economic Studies*, 74(4), 1011-1038.
- Asheim, B., Boschma, R., & Cooke, P. (2011). Constructing regional advantage: Platform polices based on related variety and differentiated knowledge bases. *Regional Studies*, 45(7), 893-904.
- Atzema, O., Lambooy, J., van Rietbergen, T., & Wever, E. (2002). *Ruimtelijke economische dynamiek; kijk op bedrijfslocatie en regionale ontwikkeling*. Bussum: Coutinho, tweede druk.
- Audretsch, D., & Fritsch, M. (1994). On the measurement of entry rates. *Empirica*, 21, 105–113.
- Audretsch, D., & Lehmann, E. (2005). Does the Knowledge Spillover Theory of Entrepreneurship hold for regions? *Research Policy*, 34, 1191–1202.
- Baimai, B. (2017). Startup: Definition, Importance, and Research Guidelines. *Executive Journal*, 37(2), 10-21.
- Bartelsman, E., & Doms, M. (2000). Understanding productivity: Lessons from longitudinal microdata. *Journal of Economic literature*, 38(3), 569-594.
- Bathelt, H., Malmberg, A., & Maskell, P. (2004). Clusters and knowledge, local buzz, global pipelines and the process of knowledge creation. *Progress in Human Geography*, 28(1), 31-56.
- Baumol, W. (1996). Entrepreneurship: Productive, unproductive, and destructive. *Journal of business venturing*, 11(1), 3-22.
- Baumol, W. (2010). *The Microtheory of Innovative Entrepreneurship*. Princeton, NJ: Princeton University Press.
- Beaudry, C., Héroux-Vaillancourt, M., & Rietsch, C. (2016). Validation of a web mining technique to measure innovation in high technology Canadian industries. *OECD Blue Sky Forum on Science and Innovation Indicators*.
- Bell, D., & Jayne, M. (2010). The creative countryside: Policy and practice in the UK rural cultural economy. *Journal of rural studies*, 26(3), 209-218.
- Birch, K., & Whittam, G. (2008). The Third Sector and the regional development of social capital. *Regional Studies*,(42), 437–450.
- Bivand, R., & Piras, G. (2015). Comparing implementations of estimation methods for spatial econometrics. *American Statistical Association*.
- Bivand, R., & Wong, D. (2018). Comparing implementations of global and local indicators of spatial association. *TEST*, 27(3), 716-748.
- Bivand, R., Pebesma, E., Gómez-Rubio, V., & Pebesma,

- E. (2008). *Applied spatial data analysis with R* (Vol. 747248717). New York: Springer.
- Bjørnstad, O., & Falck, W. (2001). Nonparametric spatial covariance functions: estimation and testing. *Environmental and Ecological Statistics*, 8(1), 53-70.
- Bjørnstad, O., Ims, R., & Lambin, X. (1999). Spatial population dynamics: Analysing patterns and processes of population synchrony, 14(11), 427-431.
- Blank, S. (2010, april 12). "Why startups are agile and opportunistic – pivoting the business model". Retrieved 2018, from [steveblank.com](https://steveblank.com/2010/04/12/why-startups-are-agile-and-opportunistic-%E2%80%93-pivoting-the-business-model/): <https://steveblank.com/2010/04/12/why-startups-are-agile-and-opportunistic-%E2%80%93-pivoting-the-business-model/>
- Blazquez, D., & Domenech, J. (2017). Is web data capable of detecting firms' activity status? Department of Economics and Social Sciences, Universitat Politècnica de València, Valencia (Spain).
- Blazquez, D., & Domenech, J. (2018). Big Data sources and methods for social and economic analyses. *Technological Forecasting and Social Change*, 130, 99-113.
- Blevins, D., & Ragozzino, R. (2018). An examination of the effects of venture capitalists on the alliance formation activity of entrepreneurial firms. *Strategic Management Journal*, 39(7), 2075-2091.
- Block, F. (2008). Swimming against the current: the rise of a hidden developmental state in the United States. *Politics and Society*, 36, 169-206.
- Boeing, G., & Waddell, P. (2016). New Insights into Rental Housing Markets across the United States: Web Scraping and Analyzing Craigslist Rental Listings. *Journal of Planning Education and Research*, 37(4), 457-476.
- Bogers, M. (2011). The open innovation paradox: knowledge sharing and protection in R&D collaborations. *European Journal of Innovation Management*, 14(1), 93-117.
- Bonnema, M. (2017). Uit de schaduw van Amsterdam: een vergelijking van het aantal innovatieve startups tussen Utrecht en Amsterdam. Universiteit Utrecht.
- Boschma, R., & Frenken, K. (2009). Technological relatedness and regional branching. In R. Boschma, K. Frenken, H. Bathelt, & M. Feldman, *Beyond territory: Dynamic geographies of knowledge creation, diffusion and innovation*, 64-81. Routledge, Taylor and Francis .
- Boschma, R., & Lammarino, S. (2009). Related variety, trade linkages, and regional growth in Italy. *Economic geography*, 85(3), 289-311.
- Boschma, R., & Wenting, R. (2007). The spatial evolution of the British automobile industry. Does location matter? *Industrial and Corporate Change*, 16(2), 213–238.
- Boschma, R., Balland, P., & Kogler, D. (2014). Relatedness and technological change in cities: the rise and fall of technological knowledge in US metropolitan areas from 1981 to 2010. *Industrial and corporate change*, 24(1), 223-250.
- Bosma, N., & Wennekers, S. (2011). Intrapreneurship versus independent entrepreneurship: A cross-national analysis of individual entrepreneurial behavior. Discussion Paper Series/Tjalling C. Koopmans Research Institute, 11(4).
- Braunerhjelm, P., Ding, D., & Thulin, P. (2018). The knowledge spillover theory of intrapreneurship. *Small business economics*, 51(1), 1-30.
- Broekel, T., & Brachert, M. (2015). The structure and evolution of inter-sectoral technological complementarity in R&D in Germany from 1990 to 2011. *Journal of evolutionary economics*, 25(4), 755-785.
- Brown, R. (2016). Mission impossible? Entrepreneurial universities and peripheral regional innovation systems. *Industry and Innovation*, 23(2), 189–205.
- Brown, R., & Mason, C. (2017). Looking inside the spiky bits: a critical review and conceptualisation of entrepreneurial ecosystems. *Small Business Economics*, 49(1), 11-30.
- Buczkowska, S., & De Lapparent, M. (2014). Location choices of newly created establishments: Spatial patterns at the aggregate level. *Regional Science and Urban Economics*, 48, 68-81.
- Camelo-Ordaz, C., Fernández-Alles, M., Ruiz-Navarro, J., & Sousa-Ginel, E. (2012). The intrapreneur and innovation in creative firms. *International Small Business Journal*, 30(5), 513-535.
- Camelo-Ordaz, C., Fernández-Alles, M., Ruiz-Navarro, J., & Sousa-Ginel, E. (2012). The intrapreneur and innovation in creative firms. *International Small Business Journal*, 30(5), 513-535.
- Castaldi, C., Frenken, K., & Los, B. (2014). Related variety, unrelated variety and technological breakthroughs: An analysis of U.S. state-level patenting. *Regional Studies*, 13(3).
- Castellacci, F., & Mee Lie, C. (2015). Do the Effects of R&D Tax Credits Vary Across Industries? A Meta-regression Analysis. *Research Policy*, 44(4), 819–32.
- Centraal Bureau voor de Statistiek. (2014). Monitor topsectoren 2014 - uitkomsten 2010, 2011 en 2012.
- Centraal Bureau voor de Statistiek. (2015). Bedrijven met website, 2012-2015 [data file]. Retrieved from [cbs.nl](https://www.cbs.nl/-/media/imported/documents/2015/44/2015__website_van_bedrijven_2012-2015_maatwerk.xls): https://www.cbs.nl/-/media/imported/documents/2015/44/2015__website_van_bedrijven_2012-2015_maatwerk.xls
- Centraal Bureau voor de Statistiek. (2016). Documentatierapport Community Innovation Survey (CIS) 2014. Centraal Bureau voor de Statistiek.
- Centraal Bureau voor de Statistiek. (2017). Monitor

- topsectoren 2017 - Methodebeschrijving en tabellenset. Den Haag.
- Centraal Bureau voor de Statistiek. (2018). ICT-gebruik bij kleine bedrijven; bedrijfsgrootte, 2017. Retrieved 2018, from [opendata.cbs.nl/statline: https://opendata.cbs.nl/statline/#/CBS/nl/dataset/83954NED/table?ts=1533726964875](https://opendata.cbs.nl/statline/#/CBS/nl/dataset/83954NED/table?ts=1533726964875)
- Centraal Bureau voor de Statistiek. (2018, december). Innovatie bij kleine bedrijven. Retrieved 2019, from [opendata.cbs.nl/statline: https://opendata.cbs.nl/statline/#/CBS/nl/dataset/84410NED/table?ts=1548676565852](https://opendata.cbs.nl/statline/#/CBS/nl/dataset/84410NED/table?ts=1548676565852)
- Centraal Bureau voor de Statistiek. (zd). Innovatie statistiek. Retrieved 2018, from [cbs.nl: https://www.cbs.nl/nl-nl/onze-diensten/methoden/onderzoeksomschrijvingen/korte-onderzoeksbeschrijvingen/innovatie-statistiek](https://www.cbs.nl/nl-nl/onze-diensten/methoden/onderzoeksomschrijvingen/korte-onderzoeksbeschrijvingen/innovatie-statistiek)
- Chen, H., Chiang, R., & Storey, V. (2012). Business intelligence and analytics: from big data to big impact. *MIS quarterly*, 1165-1188.
- Cho, C., Poudyal, N., & Roberts, R. (2008). Spatial analysis of the amenity value of green open space. *Ecological Economics*, 66(2), 403-416.
- Christaller, W. (1966). *Central places in southern Germany*. Prentice Hall.
- Coenen, L., Asheim, B., Bugge, M., & Herstad, S. (2017). Advancing regional innovation systems: What does evolutionary economic geography bring to the policy table? *Environment and Planning C: Politics and Space*, 34(4), 600-620.
- Colombo, L., Dawid, H., Piva, M., & Vivarelli, M. (2017). Does easy start-up formation hamper incumbents' R&D investment? *Small Business Economics*, 49(3), 513-531.
- Colombo, M., & Piva, E. (2008). Strengths and weaknesses of academic start-ups: a conceptual model. *IEEE Transactions on Engineering Management*, 55(1), 37-49.
- Copus, A., Courtney, P., Dax, T., Meredith, D., Naguera, J., Shucksmith, M., & Talbot, H. (2011). *Applied Research 2013/1/2 Final Report*. EDORA: European Development Opportunities for Rural Areas. Luxembourg: ESPON & UHI Millennium Institute.
- CrunchBase. (n.d.). Global dataset on funding, exits, and locations of startups and investors. Retrieved 2018, from <https://www.crunchbase.com/>
- Cukier, D., Kon, F., & Lyons, T. (2016). Software startup ecosystems evolution: The New York City case study. *Proceedings 2nd International Workshop on Software Startups*.
- Dahl, M., & Reichstein, T. (2007). Are you experienced? Prior experience and the survival of new organizations. *Industry and Innovation*, 14, 497-511.
- Davila, A., Foster, G., & Gupta, M. (2003). Venture capital financing and the growth of start-up firms. *Journal of Business Venturing*, 18(6), 689-709.
- De Groot, H., Poot, J., & Smit, M. (2007). Agglomeration, innovation and regional development. 79(3).
- Dealroom. (2018). Global dataset on funding, exits, and locations of startups and investors. Retrieved from <https://dealroom.co/>
- Deller, S., Tsai, T., Marcouiller, D., & English, D. (2001). The role of amenities and quality of life in rural economic growth. *American Journal of Agricultural Economics*, 83(3), 352-365.
- Desrochers, P., & Hospers, G. (2007). Cities and the Economic Development of Nations: An Essay on Jane Jacobs' Contribution to Economic Theory. *Canadian Journal of Regional Science*, 30(1).
- Desrochers, P., & Sautet, F. (2008). Entrepreneurial policy: The case of regional specialization vs. spontaneous industrial diversity. *Entrepreneurship Theory and Practice*, 32(5), 813-832.
- Dienst Uitvoering Onderwijs (DUO). (2019). Instellingen HBO en WO. Retrieved from [duo.nl: https://duo.nl/open Onderwijsdata/databestanden/ho/adressen/adressen-ho1.jsp](https://duo.nl/open Onderwijsdata/databestanden/ho/adressen/adressen-ho1.jsp)
- Dijkstra, L., & Ruiz, V. (2010). Refinement of the OECD regional typology: Economic performance of remote rural regions. *European Regional Science Association*.
- Dormann, F., McPherson, C., Araújo, J., Bivand, R., Bolliger, J., Carl, G., . . . Wilson, R. (2007). Methods to account for spatial autocorrelation in the analysis of species distributional data: a review. *Ecography*, 30(5), 609-628.
- EBU Research. (2016). *Nulmeting startup ecosysteem regio Utrecht*.
- Egeln, J., Gottschalk, S., & Rammer, C. (2004). Location decisions of spin-offs from public research institutions. *Industry and Innovation*, 11(3), 207-223.
- Eliasson, G. (2000). Industrial policy, competence blocs and the role of science in economic development. *Journal of Evolutionary Economics*, 10(1-2), 217-241.
- Enkel, E., & Gassmann, O. (2010). Creative imitation: exploring the case of cross-industry innovation. *R&D Management*, 30(3), 256-270.
- Epperson, B. (1993). Recent advances in correlation studies of spatial patterns of genetic variation (pp. 95-155). In M. Hecht, R. MacIntyre R.J., & M. Clegg, *Evolutionary Biology (Vols. Evolutionary Biology, vol 27)*. Boston, MA.: Springer, .
- European Commission. (2013). *Entrepreneurship 2020 Action Plan: Reigniting the entrepreneurial spirit in Europe*.
- Eurostat. (2017). Latest Community Innovation Survey results now available. Retrieved from [ec.europa.eu: http://ec.europa.eu/eurostat/web/products-eurostat-news/-/](http://ec.europa.eu/eurostat/web/products-eurostat-news/-/)

- DDN-20170124-2?inheritRedirect=true
- Faems, D. (2017, december 12). Open innovatie is geen wondermiddel. *AD/Algemeen Dagblad*, p. 20.
- Faems, D., & Wennink, S. (2018). *Innovatie Benchmark Noord-Nederland 2018*. Rijksuniversiteit Groning en Samenwerkingsverband Noord-Nederland.
- Fairlie, R., Morelix, A., Reedy, E., & Russel-Fritch, J. (2015). *The Kauffman Index 2015: Startup Activity | National Trends*.
- Fazio, C., Guzman, J., Murray, F., & Stern, S. (2016). *A new view of the skew: A quantitative assessment of the quality of American entrepreneurship*. Cambridge, MA: MIT Laboratory for Innovation Science and Policy.
- Flor, M., & Oltra, M. (2004). Identification of innovating firms through technological innovation indicators: an application to the Spanish ceramic tile industry. *Research Policy*, 33(2), 323-336.
- Florida, R. (2002). *The rise of the creative class (Vol. 9)*. New York: Basic books.
- Florida, R., & King, K. (2016). *Rise of the Urban Startup Neighborhood: Mapping Micro-Clusters of Venture Capital-Based Startups*.
- Frenken, K. (2011). Kenniseconomie in evolutionair perspectief. *TPEdigitaal*, 5(1), 97-108.
- Frenken, K., Van Oort, F., & Verburg, T. (2007). Related variety, unrelated variety and regional economic growth. *Regional Studies*, 41(5), 685-697.
- Gök, A., Waterworth, A., & Shapira, P. (2015). Use of web mining in studying innovation. *Scientometrics*, 201(1), 653-671.
- Gandomi, A., & Haider, M. (2015). Beyond the hype: Big data concepts, methods, and analytics. *International Journal of Information Management*, 35(2), 137-144.
- Gapp, R., & Fisher, R. (2007). Developing an intrapreneur-led three-phase model of innovation. *International Journal of Entrepreneurial Behavior & Research*, 13(6), 330-348.
- Garcia, R., & Calantone, R. (2002). A critical look at technological innovation typology and innovativeness terminology: a literature review. *Journal of product innovation management*, 19(2), 110-132.
- Gauthier, J., Penzel, M., & Marmer, M. (2017). *Global startup ecosystem report 2017*. San Francisco: Startup Genome.
- Gauthier, J., Stangler, D., Penzel, M., & Morelix, A. (2018). *Global Startup Ecosystem Report 2018*. San Francisco: Startup Genome.
- Glaeser, E. (2011). *Triumph of the City*. New York: Penguin.
- Glaeser, E., Kallal, H., Scheinkman, J., & Shleifer, A. (1992). Growth in cities. *Journal of political economy*, 100(6), 1126-1152.
- Godin, B. (2002). The rise of innovation surveys: Measuring a fuzzy concept. *Canadian Science and Innovation Indicators Consortium, Project on the History and Sociology of S&T Statistics, Paper*, 16.
- Godin, B. (2008). *Innovation: the History of a Category*. Project on the Intellectual History of Innovation Working Paper. (1), 1-67.
- Graham, P. (2012). *Startup = growth*. Zelf gepubliceerd online artikel. Retrieved 2018, from paulgraham.com: <http://www.paulgraham.com/growth>
- Guerrero, M., Urbano, D., Cunningham, J., & Gajon, E. (2018). Determinants of Graduates' Start-Ups Creation across a Multi-Campus Entrepreneurial University: The Case of Monterrey Institute of Technology and Higher Education. *Journal of Small Business Management*, 56(1), 150-178.
- Guerrero, M., Urbano, D., Fayolle, A., & Klofsten, M. (2016). Entrepreneurial universities: emerging models in the new social and economic landscape. *Small Business Economics*, 1-13.
- Hackl, P. (2016). Big Data: What can official statistics expect? *Statistical Journal of the IAOS*, 32(1), 43-52.
- Harrison, B. (2006). Industrial districts: old wine in new bottles?(Volume 26, number 5, 1992) . *Regional studies*, , 41(S1), S107-S121.
- Harrison, R., Mason, C., & Robson, P. (2010). Determinants of long-distance investing by business angels in the UK. *Entrepreneurship and Regional Development*, 22(2), 113-137.
- Héroux-Vaillancourt, M., & Beaudry, C. (2018, september). Validation of innovation indicators from companies' websites. 2nd International Conference on Advanced Reserach Methods and Analytics (CARMA 2018) (pp. 133-140). Editorial Universitat Politècnica de València.
- Hessisches Statistisches Landesamt. (2018). *Web scraping from company websites and machine learning for the purposes of gaining new digital data*. Working paper.
- Hisrich, R. (1990). Entrepreneurship/intrapreneurship. *American Psychologist*, 35(2), 209-222.
- Hitt, M., Ireland, R., Camp, S., & Sexton, D. (2002). Strategic entrepreneurship: Integrating entrepreneurial and strategic management perspectives. *Strategic entrepreneurship: Creating a new mindset*, 1-16.
- Hogeschool Van Hall Larenstein [VHL]. (zd). *Alle opleidingen*. Retrieved from hvhl.nl: <https://www.hvhl.nl/studeren/opleidingen/Brochure+aanvragen>
- Hollanders, H., Es-Sadki, N., & Kaner, M. (2016). *Regional Innovation Scoreboard 2016 Methodology report*. European Commission.
- Isenberg, D. (2011). *The entrepreneurship ecosystem strategy as a new paradigm for economic policy: Principles for cultivating entrepreneurship*. Presentation at the Institute of International and European Affairs.

- Jacobs, D. (2000). *Industriebeleid in de kenniseconomie. IWT-Vlaanderen, Clusterbeleid als hefboom tot innovatie*. IWT.
- Jacobs, J. (1961). *The death and life of great American cities*. New-York, NY: Vintage.
- Jacobs, J. (1970). *The Economy of Cities*. New York: Vintage.
- Janssen, M. (2015). Cross-specialization and structural holes: the case of the Dutch Topsectors. *Papers in Evolutionary Economic Geography*, 15, 19.
- Jarmin, R., & Miranda, J. (2002). *The Longitudinal Business Database, Working Paper CES-02-17*.
- Jennissen, R. (2011). *De Nederlandse migratiekaart - Achtergronden en ontwikkelingen van verschillende internationale migratietypen*. Centraal Bureau voor de Statistiek (CBS) en Wetenschappelijk Onderzoek- en Documentatiecentrum, Ministerie van Veiligheid en Justitie. Boom Juridische uitgevers.
- Johansson, B., Johansson, S., & Wallin, T. (2015). Internal and external knowledge and introduction of export varieties. *The World Economy*, 38(4), 629-654.
- Johnson, P. (1978). *New Firms and Regional Development: Some Issues and Evidence*.
- Jung, Y. Y.-b.-4. (2008, September). A web-based automated system for industry and occupation coding. In *proceedings of the International Conference on Web Information Systems Engineering* (pp. 443-457). Berlin: Springer.
- Kahle, B. (2018). Wayback Machine. Retrieved 2018, from archive.org/web/: <http://archive.org/web/>.
- Kamer van Koophandel (KvK). (2017). *Rechtsvormen - Kiezen voor de juiste rechtsvorm*. Retrieved from kvk.nl: https://www.kvk.nl/download/Rechtsvormen_01_KvK_Factsheet201511_tcm109-365627.pdf
- Kamer van Koophandel (KvK). (2018). *KvK Zzp-dynamiek - Jaaroverzicht 2017*.
- Kamer van Koophandel (KvK). (2019). *Hoe beperk je ongewenste (commerciële) benadering?* Retrieved april 2019, from kvk.nl: <https://www.kvk.nl/inschrijven-en-wijzigen/inschrijven-bij-de-kamer-van-koophandel/jouw-gegevens-in-het-handelsregister/hoe-beperk-je-ongewenste-commerciële-benadering/>
- Kamer van Koophandel. (zd). *Moet ik mijn bedrijf inschrijven?* Retrieved 2018, from kvk.nl: <https://www.kvk.nl/inschrijven-en-wijzigen/inschrijven-bij-de-kamer-van-koophandel/moet-ik-mijn-bedrijf-inschrijven/criteria-van-een-onderneming/>
- Katila, R., Chen, E., & Piezunka, H. (2012). All the right moves: How entrepreneurial firms compete effectively. *Strategic Entrepreneurship Journal*, Strategic Entrepreneurship, 6(2), 116-132.
- Kaufmann, D., & Schwartz, D. (2008). *Networking: the 'Missing Link' in public R&D support schemes*. *European Planning Studies*, 16(3), 429-440.
- Kelejian, H., & Prucha, I. (2010). Specification and estimation of spatial autoregressive models with autoregressive and heteroskedastic disturbances. *Journal of Econometrics*, 157, 53-67.
- Kinne, J., & Resch, B. (2017). Analyzing and Predicting Micro-Location Patterns of Software Firms. *Journal of Geo-Information*, 7(1), 1.
- Kinne, J., & Resch, B. (2018). Generating Big Spatial Data on Firm Innovation Activity from Text-Mined Firm Websites. *GI_Forum*(1), 82-89.
- Kirchhoff, B. (1994). *Entrepreneurship and dynamic capitalism: The economics of business firm formation and growth*. ABC-CLIO.
- Kitchin, R. (2013). Big data and human geography: Opportunities, challenges and risks. *Dialogues in human geography*, 3(3), 262-267.
- Kleiber, C., & Zeileis, A. (2016). Visualizing count data regressions using rootograms. *The American Statistician*, 70(3), 296-303.
- Kleinknecht, A., Van Montfort, K., & Brouwer, E. (2002). The non-trivial choice between innovation indicators. *Economics of Innovation and New Technology*, 11(2), 109-121.
- Klepper, S. (2002). The capabilities of new firms and the evolution of the U.S. automobile industry. *Industrial and Corporate Change*, 11, 645-666.
- Kloek, W. (2002). *Demografie van bedrijven: Uitgebreide toelichting*. Voorburg: CBS.
- Kwon, O., Lee, N., & Shin, B. (2014). Data quality management, data usage experience and acquisition intention of big data analytics *International Journal of Information Management*. 34(3), 387-394.
- Lambert, D., Brown, J., & Florax, R. (2010). A two-step estimator for a spatial lag model of counts: Theory, small sample performance and an application. *Regional Science and Urban Economics*, 40(4), 241-252.
- Larédo, P., Köhler, C., & Rammer, C. (2016). The Impact of Fiscal Incentives for R&D. In J. Edler, P. Cunningham, A. Gök, & P. Shapira (eds), *Handbook of Innovation Policy Impact*. Cheltenham: Edward Elgar Publishing, 18-53.
- Leong, C., Tan, B., Xiao, X., Tan, F., & Sun, Y. (2017). Nurturing a FinTech ecosystem: The case of a youth microloan startup in China. *International Journal of Information Management*, 37(2), 92-97.
- LeSage, J., & Pace, R. (2014). The biggest myth in spatial econometrics. *Econometrics*, 2(4), 217-249.
- Li, H., Calder, C., & Cressie, N. (2007). Beyond Moran's I: testing for spatial dependence based on the spatial autoregressive model. *Geographical Analysis*, 39(4), 357-375.
- Li, Y., Arora, S., Youtie, J., & Shapira, P. (2016). Using web

- mining to explore Triple Helix influences on growth in small and mid-size firms. *Technovation*.
- Lipi, R. (2013). The Growth Performance of Small Business under the View of Life-Cycle Model. *Management Research and Practice*, 5(4), 58-67.
- Liviano, D., & Arauzo-Carod, J. (2013). Industrial location and interpretation of zero counts. *The Annals of Regional Science*, 50(2), 515-534.
- Louw, E., Needham, D., Holden, H., & Pen, C. (2004). *Planning van bedrijventerreinen*. Den Haag: Sdu uitgevers.
- Lucas, D., & Fuller, C. (2017). Entrepreneurship: Productive, unproductive, and destructive—Relative to what? *Journal of Business Venturing Insights*. 7, 45-49.
- Luger, M., & Koo, J. (2005). Defining and tracking business start-ups. *Small Business Economics*, 24(1), 17-28.
- Möller, K. (2014). Culturally clustered or in the cloud? Location of internet start-ups in Berlin (SERC Discussion Paper No. 157). SERC Discussion Paper, 157.
- Malecki, E. (2011). Connecting local entrepreneurial ecosystems to global innovation networks: Open innovation, double networks and knowledge integration. *International Journal of Entrepreneurship and Innovation Management*, 14, 36–59.
- Malecki, E. (2018). Entrepreneurship and entrepreneurial ecosystems. *Geography Compass*.
- Malizia, E., & Motoyama, Y. (2018). Vibrant Centers as Locations for High-Growth Firms: An Analysis of Thirty US Metropolitan Areas. *The Professional Geographer*, 1-14.
- Manning, C., Raghavan, P., & Schültze, H. (2009). *Introduction to Information Retrieval*. USA: Cambridge University Press.
- Markeson, B., & Deller, S. (2012). Growth of Rural US Non-Farm Proprietors with a Focus on Amenities. *Review of Urban & Regional Development Studies*, 24(3), 83-105.
- Marsan, G., & Maguire, K. (2011). Categorisation of OECD regions using innovation-related variables. *Organisatie voor Economische Samenwerking en Ontwikkeling (OESO of OECD in het engels)*.
- Marshall, A. (1890). *Principles of economics*. London: Macmillan.
- Mason, C. (1983). Some definitional difficulties in new firms research. *Area*, 15(1), 53-60.
- Mason, C., & Brown, R. (2014). Entrepreneurial ecosystems and growth oriented entrepreneurship. *Final Report to OECD*, 30(1), 77-102.
- Mason, C., & Harrison, R. (2006). After the exit: Acquisitions, entrepreneurial recycling and regional economic development. *Regional Studies*, 40(1), 55–73.
- Mason, C., Cooper, S., & Harrison, R. (2004). Venture capital and high technology clusters: the case of Ottawa. In R. Oahey, W. Durning, & S. Kauser (Eds.), *New Technology-Based Firms in the New Millennium* (pp. 261–278).
- Matula, D., & Sokal, R. (1980). Properties of Gabriel graphs relevant to geographic variation research and the clustering of points in the plane. *Geographic Analysis*, 12(3), 205-222.
- Mayer, H. (2013a). Entrepreneurship in a hub-and-spoke industrial district: Firm survey evidence from Seattle's technology industry. *Regional Studies*, 47(10), 1715-1733.
- Mayer, H. (2013b). Firm building and entrepreneurship in second-tier high-tech regions. *European Planning Studies*, 21, 1392–1417.
- McCann, P., & Ortega-Argilés, R. (2015). Smart specialization, regional growth and applications to European Union cohesion policy. *Regional Studies*, 49(8), 1291-1302.
- McDowell, T. (2017). *A COMPREHENSIVE STUDY OF INTRAPRENEURSHIP AS A CULTURAL FORM OF INNOVATION AND SUSTAINABLE COMPETITIVE ADVANTAGE* (Doctoral dissertation).
- McGranahan, D., & Wojan, T. (2007). Recasting the creative class to examine growth processes in rural and urban counties. *Regional studies*, 41(2), 197-216.
- McGranahan, D., Wojan, T., & Lambert, D. (2011). The rural growth trifecta: outdoor amenities, creative class and entrepreneurial context. *Journal of Economic Geography*, 11(3), 529-557.
- McHenry, J. (2011). Rural empowerment through the arts: The role of the arts in civic and social participation in the Mid West region of Western Australia. *Journal of Rural Studies*, 27(3), 245-253.
- McPhee, P. (2010). An exploratory investigation of small firm export performance using innovation indicator matrices within an input-output framework. PHD, University of Wollongon, Department of Economics.
- McQueen, D., & Wallmark, J. (1982). Spin-off Companies from Chalmers University of Technology. *Technovation*, 1, 305-315.
- Mennis, J., & Guo, D. (2009). Spatial data mining and geographic knowledge discovery - An introduction. *Computers, Environment and Urban Systems*, 33(6), 403-408.
- Miles, R., Snow, C., Meyer, A., & Coleman, J. (1978). Organizational strategy, structure, and process. *Academy of management review*, 3(3), 546-562.
- Mohnen, P., Vankan, A., & Verspagen, B. (2017). Evaluating the Innovation Box Tax Policy Instrument in the Netherlands, 2007–13. *Oxford Review of Economic Policy*, 33(1), 141–56.
- Moran, P. (1950). Notes on continuous stochastic

- phenomena. *Biometrika*, 37((1/2)), 17-23.
- Morse, S. (2005). Converting Addresses to/from Latitude/Longitude/Altitude in One Step. Retrieved from [stevemorse.org: https://stevemorse.org/jcal/latlon.php](https://stevemorse.org/jcal/latlon.php)
- Mowery, D., Oxley, J., & Silverman, B. (1996). Strategic alliances and interfirm knowledge transfer. *Strategic management journal*, 17(S2), 77-91.
- Mustar, P., Wright, M., & Clarysse, B. (2008). University spin-off firms: lessons from ten years of experience in Europe. *Science and Public Policy*, 35(2), 67-80.
- Naldi, L., Nilsson, P., Westlund, H., & Wixe, S. (2015). What is smart rural development? *Journal of rural studies*, 40, 90-101.
- Neffke, F., Henning, M., & Boschma, R. (2011). How do regions diversify over time? Industry relatedness and the development of new growth paths in regions. *Economic Geography*, 87(3), 237-265.
- Nilsson, J. (1998). Introduction to Machine Learning: An early draft of a proposed Textbook [unpublished].
- Nilsson, J. (2015, februari 15). Introduction to Machine Learning. Retrieved from [robotics.stanford.edu: http://robotics.stanford.edu/people/nilsson/mlbook.html](http://robotics.stanford.edu/people/nilsson/mlbook.html)
- Nooteboom, B. (2000). *Learning and Innovation in Organizations and Economies*. Oxford: Oxford University Press.
- Nooteboom, B., Van Haverbeke, W., Duysters, G., Gilsing, V., & Van den Oord, A. (2007). Optimal cognitive distance and absorptive capacity. *Research policy*, 36(7), 1016-1034.
- Noseleit, F., de Lange, J., & Rambharos, N. (2016). Start-up activity in Groningen and the Northern Netherlands: Evolution, job creation potential, and founders' socioeconomic background. Groningen, mimeo.
- Nowlan, D. (1997). Jane Jacobs among the economist. *Ideas the Matter. The Worlds of Jane Jacobs*. 111-113.
- Oostveen, M. (2018). Protecting Individuals Against the Negative Impact of Big Data: Potential and Limitations of the Privacy and Data Protection Law Approach. Universiteit van Amsterdam (UvA). Wolters Kluwer.
- Orb Intelligence. (2018). Orb Intelligence: global dataset on company information. Retrieved from [orb-intelligence.com: http://orb-intelligence.com/](http://orb-intelligence.com/)
- Pan, F., & Yang, B. (2018). Financial development and the geographies of startup cities: evidence from China. *Small Business Economics*, 1-16.
- Parker, S. (2011). Intrapreneurship or entrepreneurship? *Journal of Business Venturing*, 26(1), 19-34.
- Partridge, M., Boarnet, M., Brakman, S., & Ottaviano, G. (2012). Introduction: Whither Spatial Econometrics? *Journal of Regional Science*, 52(2), 167-171.
- Pirnay, F., Surlémont, B., & Nlemvo, F. (2003). Toward a Typology of University Spin-Offs. *Small Business Economics*, 21(4), 355-69.
- Planbureau voor de Leefomgeving. (2017, november 16). Nederland heeft behoefte aan nieuw regionaal-economisch beleid. Retrieved from [pbl.nl/nieuws: http://www.pbl.nl/nieuws/nieuwsberichten/2017/nederland-heeft-behoefte-aan-nieuw-regionaal-economische-beleid](http://www.pbl.nl/nieuws: http://www.pbl.nl/nieuws/nieuwsberichten/2017/nederland-heeft-behoefte-aan-nieuw-regionaal-economische-beleid)
- Porter, M. (1998). Clusters and the new economics of competition. *Harvard Business Review*, 76(6), 77-90.
- Porter, M. (1998). *On competition*. Harvard, MA: Harvard Business School Press.
- Qmodus. (2017). Regio Groningen-Assen: Verborgen innovatiepotentieel - Een inventarisatie en analyse van de Topsectoren en andere innovatieve sectoren. Groningen.
- Rappert, B., Webster, A., & David, C. (1999). Making Sense of Diversity and Reluctance: Academic Industrial Relations and Intellectual Property. *Research Policy*, 28(9), 873-890.
- Renski, H. (2014). The influence of industry mix on regional new firm entry. *Regional Studies*, 48(8), 1353-1370.
- Rijksoverheid. (2016). Topsectoren Hoe & Waarom. Retrieved april 2018, from [topsectoren.nl: https://www.topsectoren.nl/topsectoren/documenten/brochures/2016/02/25/hoe-en-waarom-topsectoren](https://www.topsectoren.nl/topsectoren/documenten/brochures/2016/02/25/hoe-en-waarom-topsectoren)
- Rijksoverheid. (2018). Hoe staat Nederland ervoor. Retrieved 2019, from [bedrijvenbeleidinbeeld.nl: https://www.bedrijvenbeleidinbeeld.nl/themas/innovatie/innovatie-hoe-staat-nederland-ervoor](https://www.bedrijvenbeleidinbeeld.nl: https://www.bedrijvenbeleidinbeeld.nl/themas/innovatie/innovatie-hoe-staat-nederland-ervoor)
- Rijksoverheid. (2019). Topsectoren. Retrieved from [topsectoren.nl: https://www.topsectoren.nl/](https://www.topsectoren.nl/)
- Rijksuniversiteit Groningen [RUG]. (zd). Bacheloropleidingen per faculteit. Retrieved 2019, from [rug.nl: https://www.rug.nl/bachelors/by-faculty](https://www.rug.nl/bachelors/by-faculty: https://www.rug.nl/bachelors/by-faculty)
- Rixsen, D. (2012). ICT Roadmap ondersteunt Topsectoren. Retrieved 2018, from [netkwesties.nl: https://www.netkwesties.nl/530/ict-roadmap-ondersteunt-topsectoren.htm](https://www.netkwesties.nl: https://www.netkwesties.nl/530/ict-roadmap-ondersteunt-topsectoren.htm)
- Rodrik, D. (2008). *Industry Policy for the Twenty First Century*. Harvard University, Harvard Kennedy School (HKS); Centre for Economic Policy Research (CEPR) Working Paper 4767.
- Roelands, M., van Delden, A., & Windmeijer, D. (2017). Classifying businesses by economic activity using web-based text mining. The Hague: Statistics Netherlands.
- Roundy, P. T., Brockman, B. K., & Bradshaw, M. (2017). The resilience of entrepreneurial ecosystems. *Journal of Business Venturing Insights*(8), 99-104.
- Schumpeter, J., & Backhaus, U. (2003). The Theory of Economic Development. In B. J. (eds), *Joseph Alois Schumpeter* (pp. 61-116). *The European Heritage in Economics and the Social Sciences* (Vol. 1). MA,

- Boston: Springer.
- Schwab, K. (2017). *The Global Competitiveness Report 2017–2018*. System Initiative on Shaping the Future of Economic Progress, World Economic Forum.
- Scott, M., & Bruce, R. (1987). Five Stages of Growth in Small Business. *Long Range Planning*, 20(3), 45-52.
- Shapira, P., Gök, A., & Salehi, F. (2016). Graphene enterprise: mapping innovation and business development in a strategic emerging technology. *Journal of Nanoparticle Research*, 18(9), 269.
- Shapira, P., Gök, A., Klochikhin, E., & Sensier, M. (2014). Probing “green” industry enterprises in the UK: A new identification approach. *Technological Forecasting and Social Change*, 85, 93-104. <https://doi.org/10.1016/j.techfore.2013.10.023>.
- Simler, K. (2011). What exactly separates a startup from a medium-to-large size company? Retrieved 2018, from quora.com: <https://www.quora.com/What-exactly-separates-a-startup-from-a-medium-to-large-size-company>
- Smit, M. (2010). Interview: Science parken leiden niet altijd tot meer innovatie. (J. Brand, Interviewer)
- Smith, K. (2005). Measuring innovation.
- Sorenson, O., & Audia, P. (2000). The social structure of entrepreneurial activity: Geographic concentration of footwear production in the United States, 1940–1989. *American Journal of Sociology*, 106(2), 424-462.
- Spender, J., Corvello, V., Grimaldi, M., & Rippa, P. (2017). Startups and open innovation: a review of the literature. *European Journal of Innovation Management*, 20(1), 4-30.
- Stam, E. (2007). Why butterflies don't leave: Locational behavior of entrepreneurial firms. *Economic Geography*, 83(1), 27–50.
- Stam, E. (2014). *The Dutch entrepreneurial ecosystem*. Birch Research.
- Stam, E. (2015). Entrepreneurial ecosystems and regional policy: a sympathetic critique. *European Planning Studies*, 23(9), 1759-1769.
- Stam, E., & Saberi, S. (2017). Unicorns around the globe. Retrieved from StartupJuncture: <https://startupjuncture.com/2017/03/21/unicorns-around-globe/>
- Stam, F., & Spigel, B. (2016). Entrepreneurial ecosystems. USE Discussion paper series, 13(3).
- Stanford Encyclopedia of Philosophy. (2014). Computational Linguistics. Retrieved from plato.stanford.edu: <https://plato.stanford.edu/entries/computational-linguistics/#DocRetCluApp>
- StartupDelta. (2018). INNOVATION HUBS. Retrieved april 2018, from startupdelta.org: <https://www.startupdelta.org/about-startupdelta/innovation-hubs/>
- Stel, v., & Suddle, K. (2008). The impact of new firm formation on regional development in the Netherlands. *Small Business Economics*, 30(1), 31-47.
- Stuart, T., & Sorenson, O. (n.d.). The geography of opportunity: spatial heterogeneity in founding rates and the performance of biotechnology firms. *Research policy*, 32(2), 229-253.
- Szerb, L., Lafuente, E., Horváth, K., & Páger, B. (2018). The relevance of quantity and quality entrepreneurship for regional performance: The moderating role of the entrepreneurial ecosystem. *Regional Studies*, 1-13.
- Thornton, P., & Flynn, K. (2003). Entrepreneurship, networks, and geographies. In Z. Acs, & D. Audretsch, *Handbook of Entrepreneurship Research* (pp. 401-433) (Vol. 1, pp. 401-433). Boston, MA: Springer.
- Tiefelsdorf, M., Griffith, D., & Boots, B. (1999). A variance-stabilizing coding scheme for spatial link matrices. *Environment and Planning A*, 31, 165–180.
- Torre, A., & Rallet, A. (2005). Proximity and localization. *Regional studies*, 39(1), 47-59.
- Toussaint, G. (1980). The relative neighbourhood graph of a finite planar set. *Pattern recognition*, 12(4), 261–268.
- Toussaint, G. (2014). The sphere of influence graph: Theory and applications. *International Journal of Information Technology and Computer Science*, 14(2).
- Van de Ven, A., & Poole, M. (1990). Methods for studying innovation development in the Minnesota Innovation Research Program. *Organization science*, 1(3), 313-335.
- Van der Panne, G. (2004). Agglomeration externalities: Marshall versus Jacobs. *Journal of Evolutionary Economics*, 14(5), 593-604.
- Van der Panne, G. (2007). Issues in measuring innovation. *Scientometrics*, 71(3), 495-507.
- Van der Panne, G., & Van Beers, C. (2006). On the Marshall–Jacobs controversy: it takes two to tango. *Industrial and Corporate Change*, 15(5), 877-890.
- van Dongen, F., Jonkeren, O., & Raspe, O. (2014). TOPSECTOREN EN REGIO'S - De relatie tussen vestigingsplaatsfactoren en de concentratie van de topsectoren. Planbureau voor de Leefomgeving (PBL).
- van Loenen, B., Korthals Altes, W., Groetelaers, D., & Welle Donker, F. (2016). Ontsluiten handelsregister met open data nader belicht (Kenniscentrum open data). OTB-Onderzoek voor de gebouwde omgeving. Faculteit Bouwkunde, TU Delft.
- van Rietbergen, A., Brinkman, M., Kuypers, S., Rammos, I., & Witte, P. (2008). Evolutionaire geografie. Blijvertje of modegrill? *Geografie*, 17(8).
- Van Stel, A., Carree, M., & Thurik, R. (2005). The Effect of Entrepreneurial Activity on National Economic Growth. *Small Business Economics*, 24, 311–321. doi: 10.1007/s11187-005-1996-6.
- Veenker, S., Van der Sijde, P., During, W., & Nijhof, A.

- (2004). The perception of corporate entrepreneurship in dutch organisations. Paper presented at the High-Technology Small Firms Conference. Enschede.
- Vega, S., & Elhorst, J. (2013). On spatial econometric models, spillover effects, and W. 53rd ERSA Congress. Palermo, Italy.
- Venkataraman, S. (2004). Regional transformation through technological entrepreneurship. *Journal of Business Venturing*, 19(1), 153–167.
- Verhagen, M. (2011, februari). 'Naar de Top: de hoofdlijnen van het nieuwe bedrijfslevenbeleid'. Retrieved 2018, from Nr. 1 Brief van de minister van economische zaken, landbouw en innovatie: <https://zoek.officielebekendmakingen.nl/kst-32637-1.html>
- Ward, J., & Barker, A. (2013). Undefined by data: a survey of big data definitions. University of St Andrews, School of Computer Science. arXiv preprint arXiv:1309.5821.
- Warwick, K. (2013). Beyond Industrial Policy: emerging issues and new trends. OECD Science, Technology and Industry Policy Papers No. 2.
- Westerholt, R., Resch, B., & Zipf, A. (2015). A local scale-sensitive indicator of spatial autocorrelation for assessing high-and low-value clusters in multiscale datasets. *International Journal of Geographical Information Science*, 29(5), 868-887.
- Wymer, S., & Regan, E. (2005). Factors influencing e-commerce adoption and use by small and medium businesses. *Electronic Markets*, 15(4), 438-453.
- Xiao, J., Boschma, R., & Andersson, M. (2018). Industrial diversification in Europe: The differentiated role of relatedness. *Economic Geography*, 1-36.
- York, J., & Lenox, M. (2014). Exploring the sociocultural determinants of de novo versus de alio entry in emerging industries. *Strategic Management Journal*, 35(13), 1930-1951.
- Youtie, J., Hicks, D., Shapira, P., & Horsley, P. (2012). Pathways from discovery to commercialisation: Using web sources to track small and medium-sized enterprise strategies in emerging nanotechnologies. *Technology Analysis & Strategic Management*, 24(10), 981–995. doi:10.1080/09537325.2012. 724163.
- Zeemeijer, I. (2016). Kritiek op innovatieclusters splijt de sector. *Het Financieele Dagblad*.
- Zeileis, A., Kleiber, C., & Jackman, S. (2008). Regression models for count data in R. *Journal of statistical software*, 27(8), 1-25.
- Zook, M., Shelton, T., & Poorthuis, A. (2017). Big Data and the City. In E. Elgar, draft chapter for *Handbook of Urban Geography* edited by Ronald van Kempen and Tim Schwanen. London.

B1 Steekproef-Controle

Web ID	Top25	TS-expertise	Innovatief	Startup-kenmerken	Werknemers	leeftijd o.b.v. web/LinkedIn	Wayback machine	TS-expertise	Innovatief	Startup	Validiteit-score	Type nieuwkomer
307	9	Deels;9/9;Faciliterend/ Generiek	Onbevestigd	Onbevestigd	<20	2013	2013	1	0	2	3	
811	8	Nee;0/8;Bemiddeling	Onbevestigd	Onbevestigd	<20	2013	2013; inmiddels offline	0	0	2	2	
925	7	Ja;7/7	Onbevestigd	Nee;"Kuijpers is een professionele technisch dienstverlener met ruim 1000 medewerkers en een omzet van 200 miljoen euro (over 2017)"&"Kuijpers is een familiebedrijf, actief sinds 1921."	20+	1921	2001	2	0	0	2	
1230	6	Deels;6/6;Faciliterend/ Generiek, afgestemd op klanten in topsectoren	Ja;claimt innovatief te zijn;"We are on a journey to Society 3.0. Change is happening faster than ever and we are keeping up. We want to give you the stories, showcases and articles to inspire you about this Society 3.0 given by people from all over the world within our network. We believe that access is more important than ownership, we don't want you just to see the story, we want you live it!"	Nee;in 32 landen 196 locaties met medewerkers...	20+	te oud	2005	1	2	0	3	
1364	6	Deels;6/6;Faciliterend/ Generiek	Ja;claimt innovatief te zijn;"We continue our quest for niche markets"	Ja;oprichting 2012 en 6 medewerkers zichtbaar op LinkedIn. Grotere partners wel al veel langer actief (spinoff)	<20	2012	2013; 2012 informeel	1	2	2	5	Incumbent-backed entrant; Torrgas is a Dutch, privately funded, clean technology company which has developed a state of the art process for the production of Syngas from torrefied biomass and subsequent Syngas upgrading.
1015	5	Ja;5/5;+1 Chemie	Ja;claimt innovatief te zijn;"We offer (...) Development of innovative solutions"	Onbevestigd	<20	2013	2013	2	2	2	6	Incumbent-backed entrant; "Microbial Analysis is part of the company Bioclear bv. Bioclear was established in 1988 (...) That is why in 2001 we started to offer our microbial analyses commercially to third parties (...) The number of analyses steadily increased and in 2012 (...) The formation of Microbial Analysis by Bioclear was a fact."
304	5	Ja;5/5	Ja;claimt innovatief te zijn;"Arox staat voor innovatie. Met onze nieuwste innovaties zorgen we dat bedrijven klaar voor de toekomst zijn. De techniek van morgen, vindt u vandaag al bij ons. Neem een kijkje op de pagina van Arox Smart Technologies en leer onze autonome 'tel-drone' en de smartglasses kennen."	Nee;op basis van website al sinds 2008 actief	<20	te oud	2008; 2005 crawled maar geen info/fout- melding	2	2	0	4	
630	5	Ja;5/5	Ja;claimt innovatief te zijn;"enerGQ levert unieke en innovatieve oplossingen aan zowel de industrie, grote zakelijke instellingen als aan het MKB (midden- en kleinbedrijf). Voor deze oplossingen van energie besparen is ons bedrijf diverse keren beloond in het verleden (Innovatie Top 100, Accenture Innovation Award, IBM smart camp)."	Nee;2009 opgericht en 3 medewerkers zichtbaar op LinkedIn	<20	2009	2011	2	2	0	4	

Web ID	Top25	TS-expertise	Innovatief	Startup-kenmerken	Werknemers	leeftijd o. b.v. web/LinkedIn	Wayback machine	TS-expertise	Innovatief	Startup	Validiteit-score	Type nieuwkomer
1233	5	Ja;5/5	Ja;claimt innovatief te zijn;"We develop and manufacture high-tech measurement instruments for field operations&"The Sensor Factory is at home in all stages of sensor technology development, from concept to field-testing to implementation and roll-out."	Ja;2015 opgericht en 2 medewerkers zichtbaar op LinkedIn	<20	2015	2017	2	2	2	6	Spinoff uit bedrijf; De oprichters hebben volgens het LinkedIn profiel sector gerelateerde ervaring in de R&D tak van gevestigde bedrijven."The Sensor Factory first came together over a decade ago, searching for new ideas in untapped sensor technologies and undiscovered applications.(...) The years following saw several of these bright ideas taken to the next level, becoming prototypes and test cases. And finally, sales to multinational companies." "The Sensor Factory is made up of a team of diverse, innovative professionals intent on bringing the laboratory to the process.(...) Growing and expanding into the future."
110	5	Nee;0/6;Groothandel	Onbevestigd	Ja	<20	2015	2015	0	0	2	2	
1095	5	Ja;4/5;;-1 AgriFood	Ja;claimt innovatief te zijn;"JONG, INNOVATIEF EN BETROUWBAAR"	Ja;2015 opgericht en 2 medewerkers zichtbaar op LinkedIn	<20	2017	2017	2	2	2	6	Diversifying entrant; LinkedInprofielen van de oprichters met werkervaringen in enkel de marketing en ICT.
92	4	Nee;0/4;Bemiddeling	Onbevestigd;slechts distributie van innovatie	Twijfel;suggestie dat het bedrijf al meer dan 15 jaar ervaring heeft	<20	nvt	2016	0	0	1	1	
626	4	Ja;4/4	Ja;claimt innovatief te zijn;Tracht een nog niet bestaand product op de markt te brengen.	Nee;Inmiddels niet meer levensvatbaar zowel te lezen op de website en linkedinpagina	nvt	nvt	2016; inmiddels offline archive. fo	2	2	0	4	
152	4	Ja;4/4	Ja;claimt innovatief te zijn;"we have developed a wide range of innovative processes and inventive methods through our highly creative think tank."	Nee;2005 opgericht	nvt	2005	2011	2	2	0	4	
840	4	Ja;3/4;;-1 Water	Ja;claimt innovatief te zijn;"In 2012 we decided to focus our Research & Development and portfolio extension on Women's Health. This resulted in several patents and patent applications in the field of prognostic assays to determine the risk and pregnancy related disorders."&"Unique and ready-to-use antibodies and complete innovative diagnostic kits developed for all currently available flow cytometers."	Nee;opgericht in 1987	nvt	te oud	1999	2	2	0	4	;"It was founded in cooperation with the leading University of Groningen, in the north of The Netherlands in 1987."
606	4	Ja;3/4;;-1 Energie	Ja;claimt innovatief te zijn;"de prijzenkast van Effectus-HR: Innovatieaward SNN" & "Super blij en dankbaar zijn wij met het winnen van de Jonge Ondernemers Prijs (JOP) van Noord Nederland. Met deze prijs spreekt de jury van de JOP unaniem het vertrouwen uit in ons;in Effectus-HR.Dankzij onze unieke werkwijze met innovatieve technieken als data-driven-recruitment en neuromarketing..."	Ja;oprichting 2014 en 13 medewerkers zichtbaar op LinkedIn	<20	2014	2014	2	2	2	6	Spinoff uit bedrijf; Oprichter heeft werkervaring bij gevestigde bedrijven in dezelfde sector

Web ID	Top25	TS-expertise	Innovatief	Startup-kenmerken	Werknemers	leeftijd o.b.v. web/LinkedIn	Wayback machine	TS-expertise	Innovatief Startup	Validiteit-score	Type nieuwkomer	
496	4	Ja;4/4	Ja;claimt innovatief te zijn;"COSMONIO's office is on the internet, since we always aim to work with the brightest minds in the world regardless of their location. Still, we need labs for our research and development activities, so we founded CORTEXLAB , HEALTHLAB and LIFELAB. Both of these labs are positioned in the heart of highly innovative campuses, where science, technology and creativity blend."	Ja;Team bestaat uit 11 medewerkers	<20	2013	2013; 2010 crawled, website had toen ander doel-einde	2	2	2	6	"Incumbent-backed entrant; ""COSMONIO is part of NVIDIA INCEPTION PROGRAM. Inception nurtures dedicated and exceptional startups who are revolutionizing industries with advances in AI and data science.""
367	4	Ja;4/4	Twijfel;Business intelligence	Ja;Team bestaat uit 10 medewerkers volgens website en 13 medewerkers zichtbaar op LinkedIn	<20	2015	2015	2	1	2	5	Spinoff uit bedrijf; De oprichters hebben volgens het LinkedIn profiel sector gerelateerde ervaring bij verschillende gevestigde bedrijven zoals ASML, Post NL en Alliander.
1392	4	Deels;4/4;Faciliterend/Generiek, delen en bundelen van online informatie uit topsectoren	Onbevestigd	Onbevestigd	<20	2013	2013	1	0	2	3	
714	4	Ja;4/4	Ja;claimt innovatief te zijn;Innovatieaward winner	Nee;2003 ingeschreven	<20	te oud	2003	2	2	0	4	;"En zo ontstond in 1999 de naam Grendel Games. In 2002 werd Grendel Games officieel ingeschreven bij de Kamer van Koophandel."
794	4	Ja;4/4	Ja;claimt innovatief te zijn;Focus vooral op toepassing van 3D print technologie voor Productontwerp Product ontwikkeling en bemiddeling	Ja;Team bestaat uit 2 medewerkers	<20	2015	2015	2	2	2	6	Spinoff uit bedrijf; Oprichters volgens LinkedIn voorheen werkzaam in dezelfde sector. Website:"Hulotech 3D printing & Engineering is een jong en flexibel bedrijf welke gerund wordt door 2 eigenaren welke beide een brede technische achtergrond hebben in ontwikkeling, productie en kwaliteits-management. We hebben ruim 15 jaar ervaring in het ontwerpen van complexe 3D modellen middels geavanceerde 3D CAD software. Door de jarenlange kennis en ervaring binnen de High-Tech maakindustrie (Medische-, defensie- en industriële sector)..."
604	4	Deels;9/9;Faciliterend/Generiek	Onbevestigd;economisch samenwerkingsverband	Nee; Samenwerkingsverband	nvt	nvt	2013	1	0	0	1	
471	4	Nee;0/4;Bemiddeling	Onbevestigd	Onbevestigd	<20	2011	2011	0	0	2	2	
859	4	Ja;4/4;Groothandel met revisie service	Onbevestigd	Nee	nvt	nvt	2007	2	0	0	2	
596	4	Ja;4/4	Ja;claimt innovatief te zijn;"Our range of products varies from NP reducing Biopellets, Easy feed, Reef pearls, Joes Juice, Coral Protect and Ocean Prime Copepods, which we develop and manufacture in the Netherlands."	Nee;Oprichting 2001 volgens LinkedIn profiel oprichter	nvt	te oud	2004	2	2	0	4	
883	4	Deels;4/4;Faciliterend/Generiek, afgestemd op klanten in topsectoren	Onbevestigd	Ja;Team bestaat uit 5 medewerkers	<20	2016	2016	1	0	2	3	

Web ID	Top25	TS-expertise	Innovatief	Startup-kenmerken	Werknemers	leeftijd o.b.v. web/LinkedIn	Wayback machine	TS-expertise Innovatief Startup	Validiteit-score	Type nieuwkomer		
654	4	Ja;3/4;;;-1 ICT	Ja;claimt innovatief te zijn;Innovatieaward winner	Nee;2005 opgericht, 18 medewerkers	<20	2005	2008	2	2	0	4	Spinoff uit bedrijf;"Patrick founded Fairtrasa in 2005 after a personal epiphany led him to devote his life to social entrepreneurship." (...) "He started his business career with the global consultancy firm Deloitte and was hired in 1999 by one of his clients, Glencore, the largest commodities trading firm in the world, where he quickly established himself as a Department Head." "Since 2005, Fairtrasa has pioneered a unique and innovative solution to these global challenges, impacting the lives of over 40,000 people, including farmers and their families. Our objective now is to scale this solution to reach millions more farmers worldwide."
270	4	Ja;4/4	Ja;claimt innovatief te zijn;"From these offices clients from all over Europe are being served the most innovative solutions in the market."	Ja;Project in 2011 opgericht en 3 medewerkers zichtbaar op LinkedIn	<20	2013	2013	2	2	2	6	Spinoff uit bedrijf; LinkedIn Profiel oprichter: "AgroSense system is a project we started to develop at Unicom Kft in 2011 based on my idea." Website: "AgroSense, an open source project by Corizon.";"Corizon is an IT company founded in 2013, based in the Netherlands and in Spain."
721	4	Deels;4/4;Faciliterend/Generiek, afgestemd op klanten in topsectoren;+1 Creatieve Industrie	Ja;claimt innovatief te zijn;"Bij Groveko volgen wij trends en ontwikkelingen in de markt op de voet en trachten deze altijd zo snel mogelijk te vertalen naar concepten en/of systemen waarmee wij onze klanten vooruit kunnen helpen. (...) Wij willen voor onze klanten naast leverancier ook een kennis- en sparring-partner zijn, dit komt mede tot uiting in onze Groveko Academy. Groveko Academy is het opleidingsinstituut waar wij onze kennis delen met leidinggevende en uitvoerende schoonmaakprofessionals."	Nee;meer dan 35 jaar bestaan	nvt	te oud	2018	1	2	0	3	;"Groveko is al meer dan 35 jaar een toonaangevende, service gerichte groothandel op het gebied van hygiëne. In 2015 en 2016 hebben wij met de overname van Ewepo Reinigingssystemen uit Zwolle en Frag Hygiëne uit Goirle een investering in de toekomst gedaan. Met het migreren van Groveko, Ewepo en Frag naar "Groveko 2.0", is er een krachtige, kennis gestuurde en service gerichte groothandel ontstaan." "Groveko levert al meer dan 80 jaar schoonmaak-, hygiënesystemen en disposable producten aan meer dan 6500 bedrijven in een groot aantal branches en sectoren."
465	3	Ja;3/3	Ja;claimt innovatief te zijn;"INNOVATIVE MARINE&ARCHITECTURAL DESIGNERS"	Ja;2014 opgericht en 2 medewerkers zichtbaar op LinkedIn (Company size 2-10 employees)	<20	2014	2016	2	2	2	6	Geen Spinoff;Althans op zowel website als LinkedIn is hierover niks te vinden.
1548	3	Ja;3/3	Ja;claimt innovatief te zijn;"Wij helpen mensen in de zorg om kwaliteit te leveren, omdat het altijd beter kan. Dit doen we door verbeterprojecten vorm te geven en de uitvoer ervan te begeleiden. Onze kennis over succesvol verbeteren delen we graag."&"Innoveren is veranderen en vernieuwen maar het wiel hoeft niet opnieuw uitgevonden te worden. &"Deze week viert Zorgverbeteraars haar 10e verjaardag. Al tien jaar helpen wij (z)organisaties om nóg beter te worden. Omdat beter per definitie anders is, vragen wij met regelmaat van mensen dat ze hun werkwijze veranderen."	Nee;2007 opgericht en 8 medewerkers zichtbaar op LinkedIn	<20	2007	2017	2	2	0	4	
1121	3	Ja;3/3	Ja;claimt innovatief te zijn;"zijn wij uitgegroeid tot een unieke totaal dienstverlener in utiliteitsbouw, inrichting en onderhoud van zakelijk onroerend goed met als specialisatie de medische sector"	Nee;"Sinds de oprichting in 2007"	<20	2007	2010	2	2	0	4	

Web ID	Top25	TS-expertise	Innovatief	Startup-kenmerken	Werknemers	leeftijd o.b.v. web/LinkedIn	Wayback machine	TS-expertise	Innovatief	Startup	Validiteit-score	Type nieuwkomer
1544	3	Ja;2/3;;;-1 Energie	Ja;claimt innovatief te zijn;"Deze bundeling van krachten maakt van Zorg-nv een unieke consultancy-organisatie voor de zorg en het sociale domein."	Ja;5 medewerkers volgens website;oprichting 2015 volgens LinkedIn	<20	2015	2016	2	2	2	6	Spinoff uit bedrijf;"Zorg-nv is een (openbare) maatschap van vijf maten (vennoten).Wij leveren allen een eigen, specifieke bijdrage aan onze gezamenlijke dienstverlening. Dat levert een bijzondere bundeling van krachten op; kennis, visie en ervaring die voortkomen uit de uitoefening van de volgende beroepen: huisarts, apotheker, fysiotherapeut, consultant, projectontwikkelaar"
1222	3	Ja;3/3	Ja;claimt innovatief te zijn;"Onze ontwerpen zijn energiezuinig, innovatief...."	Nee;"De SchreuderGroep is adviesbureau met 50 jaar ervaring." Linkeding oprichtingsjaar 2004	20+	2004	1998	2	2	0	4	
885	3	Ja;3/3	"Twijfel;Onduidelijk in hoeverre bedrijf zelf ontwikkelt/innoveert;"All our programs are based on up-to-date scientific research." & ""the owners and managing partners of 2M (...) had a vision of a non-standard recruitment agency where it wasn't about bulking CV's, but about providing on point quality and establishing sustainable, long term-relations. (...) They quickly built up a large network and were among the first who used social networks for their executive searches. Aiming to be pioneers with a new philosophy was one of the foundations upon which 2M International Recruitment was built." & ""We'll continue to grow and develop as a company, for example by investing in new technologies such as the usage of big data and artificial networking"	Nee;"2M International Recruitment was founded in 2008 by Albert Peters & Jos Zijlstra" Oprichting 2008 en 121 medewerkers zichtbaar op LinkedIn	20+	2008	2013; op basis van gevonden website (https://www.2mrecruitment.com/); oude website van inschrijving niet in gebruik volgens wayback-machine enkel in 2018 gecrawled	2	1	0	3	
755	3	Nee;0/3;False Positives door verwijzing naar vacatures in andere topsectoren	Onbevestigd	Nee	<20	2010	2010; 2003 crawled, website had toen ander doeleinde	0	0	0	0	
202	3	Ja;2/3;;;-1 Tuinbouw	Onbevestigd	Twijfel;"De consultants van Symbiotic hebben ieder meer dan 15 jaar leidinggevende ervaring binnen een diversiteit aan sectoren" "meer dan 25 jaar ervaring geeft ons het vertrouwen dat wij samen met u tot uitstekende resultaten komen"	<20	nvt	2013; 2004 crawled, website had toen ander doeleinde	2	0	1	3	
1580	3	Ja;2/3;;;-1 LS&H	Ja;claimt innovatief te zijn;"Original, emotional, visceral. These are all words that are used to describe Noisia's music and each is more than justified."	Nee;zowel wayback machine (2002) als website (1982) geven aan dat Noisial veel langer bestaat	<20	1982	2002	2	2	0	4	
1553	3	Ja;3/3	Ja;claimt innovatief te zijn;"In een continu veranderende omgeving maakt Zorgon mensen, processen en systemen beter. Dit doen wij door te analyseren, op impact te beoordelen en op te lossen in uw dagelijkse praktijk."	Nee;2012 oprichting en 38 werknemers zichtbaar op LinkedIn (bovendien blijkt er alleen een locatie in Amersfoort te zijn momenteel)	20+	2012	2012	2	2	0	4	
246	3	Ja;3/3	Ja;claimt innovatief te zijn;"Bij 8D Games wordt veel gebruik van brainstormsessies om ideeën te bedenken (...) om creatieve oplossingen te verzinnen."	Ja;oprichting 2012 en 15 zichtbaar op LinkedIn	<20	2012	2013	2	2	2	6	Diversifying entrant; Oprichter volgens linkedin voorheen met werkervaring in andere sector(en)

Web ID	Top25	TS-expertise	Innovatief	Startup-kenmerken	Werknemers	leeftijd o.b.v. web/LinkedIn	Wayback machine	TS-expertise	Innovatief	Startup	Validiteit-score	Type nieuwkomer
1300	3	Ja;3/3	Ja;claimt innovatief te zijn;Conceptontwikkeling "Al onze producten hebben gemeen dat ze unieke visuals"	Ja;Oprichting 2013 en team bestaat uit 2 medewerkers	<20	2013	2015; 2009 crawled maar geen info/foutmelding	2	2	2	6	Spinoff uit bedrijf; oprichters met sectorgerelateerd werkervaring bij gevestigde bedrijven
1551	3	Ja;3/3	"Ja;claimt innovatief te zijn;""We always keep in mind that our innovative solutions must be robust, effective, user friendly, scalable and must have a very limited integration time in daily work flows. This is key for us! Together with our inhouse development team and trusted partners we are busy with the development of new solutions and to improve current applications. Besides this we take all feedback and input given by our clients very seriously and we add new functionality to our current solutions on a regular base. Our VR solutions are being used globally. Therefore, we have assigned a network of sales consultants & highly skilled trainers to stay in close contact and being your nearby partner."" & ""In close collaboration with our professional partners we developed 2 solutions: VRiend and VR-Macy""	Ja;oprichting 2017 en 2 medewerkers zichtbaar op LinkedIn, althans website redirect naar nieuw bedrijf Horus VR (https://vr-iend.com/), oorspronkelijke bedrijf en website	<20	2017	2017; oorspronkelijke website sinds 2009 crawled	2	2	2	6	Spinoff uit bedrijf; oprichters met sectorgerelateerd werkervaring bij inmiddels gevestigde bedrijven
1436	3	Nee;0/3;False Positives opgegeven website blijkt niet online en ook niet terug te halen met waybackmachine, https://www.vigilant-marine.com/ is wel gevonden als website die waarschijnlijk voor het query resultaat heeft gezorgd maar blijkt een buitenlands bedrijf	Onbevestigd	Nee;niet meer levensvatbaar	nvt	nvt	nvt	0	0	0	0	
272	3	Ja;3/3	Ja;claimt innovatief te zijn;"AIM Media ontwikkelde de laatste jaren verschillende applicaties die onze klanten het voeren van het bedrijf een stuk aangenamer, efficiënter of zelfs mogelijk hebben gemaakt. AIM Media ontwikkeld applicaties zowel in opdracht van klanten als in eigen beheer."&"Dankzij onze creatieve geest lukt het bijna altijd de klant te ondersteunen met creatieve oplossingen, een concept en passende marketing."	Nee;oprichting 2008 en 13 medewerkers zichtbaar op LinkedIn	<20	2008	2013	2	2	0	4	;"AIM Media is opgericht in 2008 door Peter Bakker. In eerste instantie spreidde hij zijn creatieve geest tentoon naast zijn studie. Tegenwoordig doet hij dat Fulltime. Tjeuke en Peter zijn vanaf 2012 samen verder gegaan onder de vlag van AIM Media."
722	3	Ja;3/3	Ja;claimt innovatief te zijn;"Onze developers worden mega enthousiast van nieuwe ontwikkelingen. Van het adopteren van de laatste technieken en het vroeg onderzoek doen naar de niet gebaande paden. Samen vooroplopen in techniek maakt ons wie we zijn." & "Als Magento 2 Enterprise Partner weten wij veel van techniek en van het bedenken en uitwerken van kwalitatief hoogwaardige oplossingen. Al die kennis en onze ontdekkingen delen we graag met jou."	Nee;op 36 werknemers zichtbaar op LinkedIn	20+	2013	2013; 1998 crawled, website had toen ander doeleinde	2	2	0	4	

Web ID	Top25	TS-expertise	Innovatief	Startup-kenmerken	Werknemers	leeftijd o.b.v. web/LinkedIn	Wayback machine	TS-expertise	Innovatief Startup	Validiteit-score	Type nieuwkomer	
808	3	Ja;3/3	Onbevestigd	Ja	<20	2013	2013	2	0	2	4	;"2007- Softwarebedrijf MMSolutions geboren"(...)2013- Idee101 is geboren! Een nieuw concept met nieuwe visie" "In 2007 begonnen wij als MMSolutions. Vanuit zolderkamers en achtertuinen werkten wij toen onder andere aan websites en E-Commerce. Heel interessant, heel leuk, maar uiteindelijk niet waar wij het best in zijn. In 2010 daalde dat besef in en was de keus duidelijk: we gaan online software bouwen."
1172	3	Ja;3/3	Ja;claimt innovatief te zijn;"Het ontwikkelen van inventieve software-producten houdt ons in ontwikkeling, waardoor we ons blijven verbeteren..."	Nee;"AL 12 JAAR BETROKEN BIJ DE ICT-AMBITIE VAN TRANSPORTEURS EN TECHNENUTEN" LinkedIn info oprichting 2007 en 2 medewerkers zichtbaar op LinkedIn	<20	2007	2012	2	2	0	4	Spinoff uit bedrijf;"Reinder.NET. Maintenance 2012, de opvolger van de ServiceBase, is in Beta versie beschikbaar"
524	3	Ja;3/3	Twijfel;Onduidelijk of bedrijf zelf ook innoveert;"Met onze ervaring in veel verschillende branches voorzien we graag u van advies voor de aanschaf van de juiste techniek of software. deFine installeert techniek door heel Nederland."	Nee;al in 2008 opgericht volgens LinkedIn	<20	2008	2014	2	1	0	3	
1374	3	Ja;1/3;software oplossingen voor computersystemen maar geen ontwikkeling van systemen en materialen zelf;;-1 HTSM	Ja;claimt innovatief te zijn;ontwikkelt en onderhoudt nieuwe producten	Ja;oprichting 2012 met 1 medewerker zichtbaar op LinkedIn, aangezien het om een co-owner gaat zou er ook nog minimaal een 2e werknemer moeten zijn waarvan geen LinkedIn profiel zichtbaar is	<20	2012	2012	2	2	2	6	Spinoff uit onderzoeksinstelling; "As part of the Distributed Systems research group I assisted in the research, design, development and evaluation of office energy saving through component based automation." (linkedin)
1404	3	Ja;3/3	Ja;claim innovatief te zijn;"Van onze tafel ontwerpstudio. Wij werken aan gave boekenseries, hebben een webwinkel met eigen ontwerpen en schrijven met regelmaat op ons lifestyle blog!"	Nee;Niet meer levensvatbaar;"VAN ONZE TAFEL gaat stoppen! Wij gaan allebei een eigen richting in. Daarbij blijven we elkaar feedback geven, helpen we elkaar bij opdrachten waar nodig en inspireren we elkaar. Working apart together!!"	<20	2013	2013; sinds 2018 inactief	2	2	0	4	
1522	3	Ja;2/3;deels generiek faciliterend ("WtR (WillingtoRun) offers software management solutions for all areas of industries and organisations."), ontwikkelt zelf geen HTSM;;-1 HTSM	Twijfel;Business intelligence	Ja;oprichting 2013 met 2 medewerkers zichtbaar op LinkedIn	<20	2013	2014	2	1	2	5	Spinoff uit bedrijf; oprichter met sectorgerelateerd werkervaring bij gevestigde bedrijven
522	3	Deels;3/3;Uitvoering	Onbevestigd	Nee;al sinds 2002 actief volgens website	nvt	2002	2002	1	0	0	1	
1237	3	Ja;2/3;;-1 Energie	Twijfel;Twijfel;Onduidelijk in hoeverre bedrijf zelf ontwikkelt/innoveert;"De combinatie van ons uitgebreide portfolio met slimme en innovatieve oplossingen maakt het mogelijk om uw organisatie klaar te maken voor 'het nieuwe werken.'" & "Shareprint levert de producten en diensten om informatie altijd en overal beschikbaar te hebben via elk medium. De wereld van de communicatie verandert met de dag. Daarom biedt Shareprint oplossingen om bedrijven uit alle sectoren te laten inspelen op de veranderende wereld."	Ja;oprichting 2014 en 6 werknemers zichtbaar op LinkedIn	<20	2014	2016	2	1	2	5	Spinoff uit bedrijf; oprichter met sectorgerelateerd werkervaring bij gevestigde bedrijven
1538	3	Ja;3/3	Ja;claimt innovatief te zijn;	Nee;oprichting 2012 maar inmiddels al 51-200 werknemers volgens linkedin	20+	2012	2014	2	2	0	4	

Web ID	Top25	TS-expertise	Innovatief	Startup-kenmerken	Werknemers	leeftijd o.b.v. web/LinkedIn	Wayback machine	TS-expertise	Innovatief	Startup	Validiteit-score	Type nieuwkomer
214	3	Nee;0/3;Domein buitengebruik of gebruikt voor andere doeleinde	Onbevestigd	Nee;Website buiten gebruik	nvt	nvt	2001; sinds 2001 crawled, verschillende doeleinde domeinnaam regelmatig buiten gebruik	0	0	0	0	
889	3	Deels;3/3;Faciliterend/Generiek	Onbevestigd	Nee;"Kijk op het noorden signaleert al 48 jaar de economische, trendy, culturele en toeristische ontwikkelingen in Noord-Nederland. Via het magazine Kijk op het noorden, onze website en social media kanalen publiceren we alle belangrijke ontwikkelingen op economisch, maatschappelijk, cultureel en recreatief gebied in Noord-Nederland. Zakelijk nieuws, inspirerende verhalen, boeiende citaten en informatieve columns, maken Kijk op het noorden tot hét platform voor ondernemers in Noord-Nederland."	nvt	te oud	2013	1	0	0	1	
117	3	Ja;2/3;;;-1 AgriFood	Ja;claimt innovatief te zijn;"Wij zijn een innovatieve onderneming gevestigd in Roden (Drenthe)"	Nee;2010 opgericht volgens LinkedIn	<20	2010	2012	2	2	0	4	;"Na in de Food en Non-Food export gezeten te hebben is 15 jaar geleden de overstap gemaakt naar het ontwerpen en implementeren van software en het begeleiden van dergelijke trajecten. Onze achtergrond maakt Hunter-CRM zo flexibel."
116	3	Ja;4/4	Ja;claimt innovatief te zijn;Focus ligt vooral op toepassing van 3D print technologie voor Productontwerp Product ontwikkeling en bemiddeling	Ja;Team bestaat uit 2 medewerkers	<20	2015	2015	2	2	2	6	Spinoff uit bedrijf; Oprichters volgens LinkedIn voorheen werkzaam in dezelfde sector. Website:"Hulotech 3D printing & Engineering is een jong en flexibel bedrijf welke gerund wordt door 2 eigenaren welke beide een brede technische achtergrond hebben in ontwikkeling, productie en kwaliteits-management. We hebben ruim 15 jaar ervaring in het ontwerpen van complexe 3D modellen middels geavanceerde 3D CAD software. Door de jarenlange kennis en ervaring binnen de High-Tech maakindustrie (Medische-, defensie- en industriële sector)..."
72	3	Ja;3/3	Ja;claimt innovatief te zijn;"Elaborate process know-how and extensive technological experience provide us with a leading role advantage in the precise electrochemical machining technology's future development."&"Those process promotion activities, together with our extensive R&D knowledge capabilities have placed us at a leading thought leadership position within the pECM market. We recognize the responsibility of maintaining our knowledge authority role, and thus strive to keep on offering superior products and services to our current and potential clients."	Nee;2003 opgericht	nvt	te oud	2003	2	2	0	4	

Web ID	Top25	TS-expertise	Innovatief	Startup-kenmerken	Werknemers	leeftijd o.b.v. web/LinkedIn	Wayback machine	TS-expertise	Innovatief Startup	Validiteit-score	Type nieuwkomer	
830	3	Ja;3/3	Ja;claimt innovatief te zijn;Focus op nieuwe markten en onderscheiden: "As a result of the country's intense focus on mass production, China's economy grew dramatically during the early 2000s and a burgeoning market for luxury items such as after-market wheels developed quickly alongside it. With newfound disposable income at their fingertips, China's rapidly emerging middle class propelled sales of INOVIT products to a level equal to that in Europe within a mere two months."&"The fifth INOVIT operation was established in Northern Italy in 2010 (...) Italy is home to some of the biggest names the alloy wheel world has ever known and while competition was tough, we were determined to set ourselves apart from the crowd with our unique designs and expert engineering."	Nee	<20	1998	2004; 2000 crawled, maar geen info	2	2	0	4	;"January 13 1998 (...) The Start (...) Established by a sole trader early in 1998, the business commenced operation from a home office and small warehouse located in Friesland, a northern province in the Netherlands. After successfully sourcing suppliers for wheels and tires, the business began trading and a customer network that stretched across Northern Europe and the UK started to develop. Tenacity and a passion for business quickly led to a sharp rise in interest in the company's products and services and it soon became apparent that expansion was the necessary next step. With a growing customer base and great reputation, it was time to source a much bigger warehouse and take on some extra staff to cope with demand."
846	3	Ja;3/3	Ja;claimt innovatief te zijn;"Innovative Technical solutions" & "Vanuit Drachten ontwikkelt ITS innovatieve en technologisch geavanceerde oplossingen voor bedrijven die zelf onvoldoende ontwikkelcapaciteit of kennis in huis hebben."	Ja;LinkedIn:" Year founded 2012 (...) Company size 2-10 employees"	<20	2012	2012	2	2	2	6	Spinoff uit bedrijf;"With 22 years experience in defense (marines and airforce) I am an initiative and assertive person, a "quick thinker", able to analyse company operations, tactics and strategies because of my broad experience at all company levels. "
890	3	Deels;3/3;Faciliterend/Generiek	Onbevestigd	Nee;zie waybackmachine	nvt	te oud	2006	1	0	0	1	
1569	3	Deels;3/3;vooral gericht op uitvoering en informeren zonder zelf te innoveren	Onbevestigd	Onbevestigd	<20	2014	2014; 2013 crawled maar geen info/foutmelding	1	0	2	3	
1244	3	Deels;3/3;Faciliterend/Generiek	Onbevestigd	Ja;Geen LinkedIn gevonden	<20	2012	2012	1	0	2	3	
30	3	Deels;3/3;Uitvoering	Twijfel;maakt gebruik van innovatie maar lijkt zelf vooral een platform voor samenwerking en uitvoering;"Diary Campus Leeuwarden, een proefboerderij voor innovatie op het terrein van landbouw en veeteelt en onderdeel van de Universiteit van Wageningen, heeft de mestvergisting nieuw leven in geblazen. Evergreen Energy heeft in samenwerking met de University of Minnesota (VS) een innovatieve technologie ontwikkeld dat niet alleen zorgt dat de opbrengst van biogas uit biomassa hoger is, maar dat het ook cellulose / lignine volledig wordt vergist met betere kwaliteit biogas (meer methaan) tegen aanzienlijk lagere totale kosten."	Twijfel;Nieuwe samenwerking sinds 2015 van verschillende initiatiefnemers, wel aan het einde van de innovatieketen	<20	2015	2016	1	1	1	3	
676	3	Ja;3/3	Onbevestigd	Nee;2008 opgericht en Freelancer, lijkt niet gericht op groei	<20	2008	2008	2	0	0	2	
488	3	Ja;3/3	Onbevestigd	Ja	<20	2013	2013	2	0	2	4	
1572	3	Ja;3/3	Ja;claimt innovatief te zijn;"wij zijn er om het online landschap elke dag een stukje mooier en slimmer te maken. Wij bedenken en ontwikkelen strategische design producten die zich onderscheiden van de massa."	Nee;opgericht in 2004 en 23 medewerkers	20+	2004	2004	2	2	0	4	

Web ID	Top25	TS-expertise	Innovatief	Startup-kennmerken	Werknemers	leeftijd o.b.v. web/LinkedIn	Wayback machine	TS-expertise	Innovatief	Startup	Validiteit-score	Type nieuwkomer
115	3	Ja;3/3	Ja;claimt innovatief te zijn;"The designs we generate are presented in the collection of HuisVeendam and are exemplary of how design can speed up the innovation process and lead the way in sustainable change."	Ja;"In 2012 he cofounded HuisVeendam, an innovation platform and design label to promote starch technology and design thinking in product development, from raw material to end product and beyond."	<20	2017	2017	2	2	2	6	Diversifying entrant; "Tjeerd Veenhoven is 36 and is founder of Studio Tjeerd Veenhoven. Besides that he is part of several innovation platforms. Through workshops and teaching he spreads awareness on bio based materials and the power of design thinking. The PalmLeather project is an award winning and exemplary of how design can promote eco friendly materials and social responsibility. In 2012 he cofounded HuisVeendam, an innovation platform and design label to promote starch technology and design thinking in product development, from raw material to end product and beyond."
464	3	Ja;3/3	Ja;claimt innovatief te zijn;Ontwikkelt eigen game	Ja	<20	2015	2015	2	2	2	6	Geen Spinoff;Althans op zowel website als LinkedIn is hierover niks te vinden.
602	3	Ja;3/3	Onbevestigd	Nee;"Wij zijn opgericht in 1994 en werken vanuit kantoren in Zuidwolde (Dr.), Appingedam en Almere." & "Ons team bestaat momenteel uit 35 enthousiaste medewerkers."	>20+	1994	2012	2	0	0	2	
1398	3	Deels;3/3;Faciliterend/Generiek	Onbevestigd	Ja;oprichting 2012 en 4 medewerkers volgens webpagina en zichtbaar op LinkedIn	<20	2016	2016	1	0	2	3	
1189	3	Ja;3/3	Ja;Duidelijke missie om nieuwe oplossingen te bieden voor bestaande maatschappelijke tekortkomingen/behoefte: "Bij elke auto en elk huishoudelijk apparaat krijgen we een handleiding voor een optimale levensduur. Als moderne mens ontbreekt deze handleiding over voeding en leefstijl echter meestal."	Twijfel;Op LinkedIn zijn de huidige werkzaamheden uitgesplitst in 5 verschillende entiteiten waarvan de oprichting uiteenloopt van 2003 tot 2017.	<20	2003	2008	2	2	1	5	Incumbent-backed entrant
1049	3	Ja;3/3	Twijfel;Produceert (ver) nieuwe(nde) producten	Nee;"In business since 1907, the Nielsen-Massey brand has built an enviable reputation for the quality and consistency of its pure vanillas."	nvt	te oud	2001	2	1	0	3	
1104	3	Nee;0/3;Bemiddeling	Onbevestigd	Onbevestigd	<20	2016	2016	0	0	2	2	
1111	3	Ja;3/3	Ja;claimt innovatief te zijn;	Ja;"The idea behind Pectcof was created in an innovation course at Wageningen University in March 2011. In this course the founders came together for the first time, to talk about a biorefinery concept for a huge side stream in the industrial coffee chain: the coffee cherry."; Oprichting 2011 (LinkedIn)	<20	2011	2013	2	2	2	6	Spinoff uit onderzoeksinstelling; "The idea behind Pectcof was created in an innovation course at Wageningen University in March 2011" & "Since then, the multicultural team has been working together to make Pectcof a great success. With the use of the different backgrounds in chemical and electrical engineering, as well as Master degrees in Food Chemistry, Process Engineering, Biotechnology, Management and Marketing the team combines different skills and experiences to run the company. On top of this management team we have many experienced advisers and partners that strengthen and advise on all the important aspects of technology and running a company."

Web ID	Top25	TS-expertise	Innovatief	Startup-kenmerken	Werknemers	leeftijd o.b.v. web/LinkedIn	Wayback machine	TS-expertise	Innovatieve Startup	Validiteit-score	Type nieuwkomer	
1364	6	Deels;6/6;Faciliterend/ Generiek	Ja;claimt innovatief te zijn;"We continue our quest for niche markets"	Ja;oprichting 2012 en 6 medewerkers zichtbaar op LinkedIn. Grotere partners wel al veel langer actief (spinoff)	<20	2012	2013; 2012 informeel	1	2	2	5	Incumbent-backed entrant; Torrgas is a Dutch, privately funded, clean technology company which has developed a state of the art process for the production of Syngas from torrefied biomass and subsequent Syngas upgrading.
1564	1	Ja;2/1;;+1 AgriFood	Onbevestigd	Twijfel;Geen LinkedIn en contact pagina op website geeft foutmelding, mogelijk niet meer actief	<20	2012	2016	2	0	1	3	
1133	1	Ja;1/1	Onbevestigd	Nee;oprichting 2003 of eerder	<20	2003	2007; geeft oprichting in 2003 aan, inmiddels nieuwe website https://www.kwekerijpoppen.nl	2	0	0	2	
586	0	Ja;1/1	Onbevestigd	Nee;oprichting 2009 en 2 medewerkers volgens website en zichtbaar op LinkedIn	<20	2009	niet gevonden	2	0	0	2	
723	2	Ja;2/2	Ja;Ja;claimt innovatief te zijn;Het GvP werkt nauw samen met de Gezondheidsdienst voor Dieren, de Faculteit Diergeneeskunde en haar toeleveranciers voor het ontwikkelen van nieuwe producten."	Ja;Het team bestaat momenteel uit een groep van vier pluimveedierenartsen, drie dierenartsassistenten en een analist, 8 medewerkers volgens website en oprichting 2011 (linkedin)	<20	2011	2012	2	2	2	6	Spinoff uit bedrijf; Oprichter volgens LinkedIn voorheen werkzaam in dezelfde sector.
654	4	Ja;3/4;;;-1 ICT	Ja;claimt innovatief te zijn;Innovatieaward winner	Nee;2005 opgericht, 18 medewerkers	<20	2005	2008	2	2	0	4	
1498	1	Ja;1/1	Ja;claimt innovatief te zijn;"Wijkopenautos.nl is in 2012 opgericht met de overtuiging dat we de particuliere auto verkoop voor iedereen toegankelijker en makkelijker kunnen maken. Gezien het immense aanbod en het soms frustrerende proces bij tweedehands autoverkoop sites geloven wij dat het verkopen van een auto uit slechts drie stappen hoeft te bestaan. Kentekencheck, taxatie, verkoop. Dit alles doen we volledig gratis en zonder verplichtingen, binnen 24 uur."	Nee;"Met al meer dan 1.000.000 gekochte auto's door heel Europa en 27 vestigingen in Nederland zijn we marktleider op het gebied van tweedehands auto inkoop."	20+	2012	2014	2	2	0	4	
861	2	Ja;2/2	Twijfel;Freelance Platform	Twijfel;oprichting 2012 en 23 medewerkers zichtbaar op LinkedIn	21-25	2012	2013	2	1	1	4	
1464	1	Deels;1/1;Faciliterend/ Generiek	Onbevestigd	Onbevestigd	<20	2012	2012; inmiddels website inactief	1	0	2	3	
115	3	ja;3/3	Ja;claimt innovatief te zijn;"The designs we generate are presented in the collection of HuisVeendam and are exemplary of how design can speed up the innovation process and lead the way in sustainable change."	Ja;"In 2012 he cofounded HuisVeendam, an innovation platform and design label to promote starch technology and design thinking in product development, from raw material to end product and beyond."	<20	2017	2017	2	2	2	6	Diversifying entrant; "Tjeerd Veenhoven is 36 and is founder of Studio Tjeerd Veenhoven. Besides that he is part of several innovation platforms. Through workshops and teaching he spreads awareness on bio based materials and the power of design thinking. The PalmLeather project is an award winning and exemplary of how design can promote eco friendly materials and social responsibility. In 2012 he cofounded HuisVeendam, an innovation platform and design label to promote starch technology and design thinking in product development, from raw material to end product and beyond."

Web ID	Top25	TS-expertise	Innovatief	Startup-kenmerken	Werknemers	leeftijd o.b.v. web/LinkedIn	Wayback machine	TS-expertise	Innovatief	Startup	Validiteit-score	Type nieuwkomer
959	2	Ja;1/2;;;-1 AgriFood	Ja;claim innovatief te zijn;"it is forecast a massive spread of this technology in the coming several years. At the very beginning, it will be launched in the construction of new buildings, public spaces, manufacturing companies, sphere of entertainment and advertising. Light diodes are functionally adapted to unexpected places, therefore we are confident that there will be a large number of unusual and innovative solutions, which will be created together with YOU – OUR CUSTOMERS."	Nee;Hoofdkantoor in Litouwen	nvt	2010	2011	2	2	0	4	
176	1	Ja;1/1	Twijfel;Korte teksten op website, lijkt nog niet volwaardig;"RecycleNet focust zich voornamelijk op de recyclage van nettenmateriaal afkomstig uit de visserij, bouw, land- en tuinbouw en de sport. Deze materialen worden in de huidige recyclemarkt gekenmerkt als 'ingewikkeld en moeilijk verwerkbaar'"	Twijfel;Levensvatbaarheid onduidelijk op basis van website, en onvindbare aanvullende informatie op linkedin	nvt	2018	2018	2	1	1	4	
1544	3	Ja;2/3;;;-1 Energie	Ja;claimt innovatief te zijn;"Deze bundeling van krachten maakt van Zorg-nv een unieke consultancy-organisatie voor de zorg en het sociale domein."	Ja;5 medewerkers volgens website;oprichting 2015 volgens LinkedIn	<20	2015	2016	2	2	2	6	Spinoff uit bedrijf; "Zorg-nv is een (openbare) maatschap van vijf maten (vennoten).Wij leveren allen een eigen, specifieke bijdrage aan onze gezamenlijke dienstverlening. Dat levert een bijzondere bundeling van krachten op; kennis, visie en ervaring die voortkomen uit de uitoefening van de volgende beroepen: huisarts, apotheker, fysiotherapeut, consultant, projectontwikkelaar"
746	1	Ja;1/1	Onbevestigd	Ja;oprichting 2012 en 4 medewerkers zichtbaar op LinkedIn	<20	2014	2014	2	0	2	4	
812	1	Ja;2/2	"Ja;claimt innovatief te zijn;"Aan veel innovaties die in het werkveld hebben plaats gevonden heeft Ifmec een bijdrage geleverd. Ifmec heeft als kenniscentrum op het gebied van facility management in brede zin een eigen onderzoeksbureau en onderzoeksagenda. Al meer dan twintig jaar besteedt Ifmec aandacht aan onderzoek. De opgebouwde kennis wordt gebruikt om de opleidingen constant te blijven ontwikkelen.""	Nee;oprichting 1994 en 30 medewerkers zichtbaar op LinkedIn	20+	1994	1998	2	2	0	4	
931	1	Ja;3/1;;;+2 ICT + LS&H	Ja;claimt innovatief te zijn;"De wereld om ons heen verwondert ons elke dag. Onze nieuwsgierigheid brengt ons nieuwe vragen, nieuwe inzichten en ideeën. Lable Lab is onze speeltuin waarin we naar hartelust experimenteren met nieuwe concepten en technieken. Iedereen heeft toegang tot Lable Lab. Heb je een goed idee, gooi het in de groep en we gaan samen ontdekken wat er mogelijk is."	Nee;oprichting 2007 en 20 medewerkers zichtbaar op LinkedIn	20-25	2007	2012; 2008 crawled maar geen info/foutmelding	2	2	0	4	
1077	2	Ja;2/2;Facilerend/Generiek	Onbevestigd	Nee;opgericht 1969 en meer dan 1484 medewerkers zichtbaar op LinkedIn	20+	1969	2008; 1999 t/m 2007 crawled maar website had toen ander doeleinde (vakantiereizen)	2	0	0	2	

Web ID	Top25	TS-expertise	Innovatief	Startup-kenmerken	Werknemers	leeftijd o.b.v. web/LinkedIn	Wayback machine	TS-expertise	Innovatief Startup	Validiteit-score	Type nieuwkomer	
1095	5	Ja;4/5;;;-1 AgriFood	Ja;claimt innovatief te zijn;"JONG, INNOVATIEF EN BETROUWBAAR"	Ja;2015 opgericht en 2 medewerkers zichtbaar op LinkedIn	<20	2017	2017	2	2	2	6	Geen Spinoff; LinkedInprofielen van de oprichters met werkervaringen in enkel de marketing en ICT.
500	1	Deels;2/2;Faciliterend/Generiek	Onbevestigd	Nee;Crowe Global is en spinoff van Crowe Horwath International volgens LinkedIn. "Ranked as the eighth largest accounting network in the world, Crowe Global is a leader with over 200 independent accounting and advisory firms in more than 130 countries." & "For almost 100 years, Crowe has been making smart decisions for multinational clients working across borders."	20+	nvt	2018	1	0	0	1	
292	1	Nee;0/1	Onbevestigd	Nee;bestaat al minstens sinds 2010 (waybackmachine)	nvt	te oud	2000	0	0	0	0	
676	3	Ja;3/3	Onbevestigd	Nee;2008 opgericht en Freelancer, lijkt niet gericht op groei	<20	2008	2008	2	0	0	2	
195	1	Ja;1/1	Onbevestigd	Twijfel;lijkt niet langer levensvatbaar	<20	2016	2016; wensite inmiddels offline	2	0	1	3	
1273	2	Nee;0/2;Groothandel	Onbevestigd	Ja;"SP Dental Supplies is in 2014 opgericht door Daniël Schmidt en Jeroen Platje, samen hebben zij 12 jaar praktijk kennis zowel op klinisch als op technisch gebied."	<20	2014	2015	0	0	2	2	
486	2	Ja;1/2;;+1 ICT;-1 LS&H	Onbevestigd	Ja	<20	2015	2016	2	0	2	4	
630	5	Ja;5/5	Ja;claimt innovatief te zijn;"enerGQ levert unieke en innovatieve oplossingen aan zowel de industrie, grote zakelijke instellingen als aan het MKB (midden- en kleinbedrijf). Voor deze oplossingen van energie besparen is ons bedrijf diverse keren beloond in het verleden (Innovatie Top 100, Accenture Innovation Award, IBM smart camp)."	Nee;2006 opgericht en 3 medewerkers zichtbaar op LinkedIn	<20	2006	2011	2	2	0	4	
92	4	Nee;0/4;Bemiddeling	Onbevestigd;slechts distributie van innovatie maar mogelijk unieke	Twijfel;suggestie dat het bedrijf al meer dan 15 jaar ervaring heeft	<20	nvt	2016	0	0	1	1	
803	1	Ja;1/1;;+1 ICT	Onbevestigd	Nee;bestaat sinds 2003 (waybackmachine)	<20	2003	2003	2	0	0	2	
1363	1	Nee;0/1;Groothandel	Onbevestigd	Ja	<20	2013	2013	0	0	2	2	
307	9	Deels;9/9;Faciliterend/Generiek, afgestemd op klanten in topsectoren	Onbevestigd	Onbevestigd	<20	2013	2013	1	0	2	3	
488	3	Ja;3/3	Onbevestigd	Ja	<20	2013	2013	2	0	2	4	
569	1	Ja;1/1	Onbevestigd	Nee;"Digizijn heeft ruim vijftien jaar ervaring in ICT en web."Op LinkedIn staat 25 jaar.	<20	te oud	2011	2	0	0	2	
577	2	Ja;1/1;;+1 Creatieve Industrie	Twijfel;Promotieplatform	Ja;2013 opgericht en 13 medewerkers	<20	2013	2015	2	1	2	5	Geen Spinoff;"Newz Marketing & Media zag het levenslicht vijf jaar geleden toen drie vrienden hetzelfde bedachten: waarom helpen we bedrijven niet in het hele proces met hun promotie? Zo gezegd, zo gedaan."
1509	1	Ja;1/1	Onbevestigd	Twijfel;Freelancer, mogelijk geen groeiambities	<20	nvt	niet gevonden	2	0	1	3	
524	3	Ja;3/3	Twijfel;Onduidelijk of bedrijf zelf ook innoveert;"Met onze ervaring in veel verschillende branches voorzien we graag u van advies voor de aanschaf van de juiste techniek of software. deFine installeert techniek door heel Nederland."	Nee;al in 2008 opgericht volgens LinkedIn	<20	2008	2014	2	1	0	3	

Web ID	Top25	TS-expertise	Innovatief	Startup-kenmerken	Werknemers	leeftijd o.b.v. web/LinkedIn	Wayback machine	TS-expertise	Innovatief	Startup	Validiteit-score	Type nieuwkomer
859	4	Ja;4/4;Groothandel met revisie service	Onbevestigd	Nee	nvt	nvt	2007	2	0	0	2	
1392	4	Deels;4/4;Faciliterend/Generiek, delen en bundelen van online informatie uit topsectoren	Onbevestigd	Onbevestigd	<20	2013	2013	1	0	2	3	
1553	3	Ja;3/3	Ja;claimt innovatief te zijn;"In een continu veranderende omgeving maakt Zorgon mensen, processen en systemen beter. Dit doen wij door te analyseren, op impact te beoordelen en op te lossen in uw dagelijkse praktijk."	Nee;2012 oprichting en 38 werknemers zichtbaar op LinkedIn (bovendien blijkt er alleen een locatie in Amersfoort te zijn momenteel)	20+	2012	2012	2	2	0	4	
22	1	Nee;0/1;;+1 Logistiek	Onbevestigd	Nee;2 medewerkers maar oprichting al langer dan 15 jaar geleden	<20	te oud	2013	0	0	0	0	
112	2	Ja;1/2;;-1 Energie	Onbevestigd	Nee;1 medewerker, "In 2005 is Home-Unique Binnenhuisarchitectuur opgericht door Lisette Uitendaal."	<20	2005	2008	2	0	0	2	
575	1	Deels;1/1;Faciliterend/Generiek	Onbevestigd	Ja;oprichting 2013 en team bestaat uit 7 medewerkers	<20	2013	2016	1	0	2	3	
605	1	Ja;1/1;+1 Creatieve industrie	Twijfel;ontwerp van energie neutrale woningen	Nee;oprichting 2005 (linkedin)	<20	2015	2008 maar geen info	2	1	0	3	
854	2	Nee;0/2;Groothandel	Onbevestigd	Nee;bestaat al minstens sinds 2000 (waybackmachine)	20+	te oud	2000	0	0	0	0	
251	1	Ja;1/1	Ja;claimt innovatief te zijn;"Naast voldoende inhoudelijke kennis zijn creativiteit, bezieling en innovatiekracht belangrijke eigenschappen voor een ontwikkelaar om de steeds nieuwe vraagstukken in een veranderende omgeving het hoofd te kunnen bieden."	Nee;oprichting in 1949	20+	1949	1998	2	2	0	4	
501	1	Ja;1/1;+1 ICT	Ja;claimt innovatief te zijn;"Altijd op zoek naar nieuwe ontwikkelingen en de beste oplossingen."&"CTNET (...) ontwerpt."	Nee;oprichting 2003 en op linkedin al meer da 20 medewerkers	20+	2003	2010	2	2	0	4	
527	2	Ja;2/2	Onbevestigd	Nee;oprichting 2009	<20	2009	2012	2	0	0	2	
551	1	Ja;1/1	Ja;claimt innovatief te zijn;"Zoals het een goede Tafelheer betaamt, zit deze niet lang stil aan tafel en ontwerpt, creëert en produceert hij constant verder. Het zal dan ook niet bij de huidige ontwerpen blijven. Mocht uw droomtafel er nu nog niet tussen zitten, neem dan gerust contact met mij op en laat al uw wensen de vrije loop. Gezamenlijk kunnen we tot de mooiste eindproducten komen."	Ja;oprichting 2013 en 2 medewerkers	<20	2013	2016	2	2	2	6	Geen Spinoff; "De oorsprong van de Tafelheren ligt in Zuidlaren waar twee oud HTS studenten in 2013 het roer hebben omgegooid." Oprichter voorheen International projectmanager (technische bedrijfskunde)
760	1	Ja;1/1	Ja;claimt innovatief te zijn;"Ik ben sterk in het bedenken en uitvoeren van creatieve concepten. Of het nu gaat om grafisch ontwerp voor drukwerk of voor webapplicaties."	Ja;oprichting 2018, 1 medewerker	<20	2018	2018	2	2	2	6	Geen Spinoff; oprichter heeft volgens linkedin geen werkervaring bij gevestigde bedrijven
1134	1	Nee;0/1;Groothandel	Onbevestigd	Nee;oprichting 1999	nvt	1999	2001	0	0	0	0	
1310	2	Ja;2/2	Ja;claimt innovatief te zijn;"Taal&Digitaal staat voor creatieve internet-technologie. Daarbij gaan we grote uitdagingen niet uit de weg. Alles met het doel om de digitale wereld te verrijken"	Twijfel;zelfde eigenaar al in 2006 maar andere activiteiten, mogelijk sinds 2012 diversifying entrant, inmiddels 8 werknemers	<20	2006	2006	2	2	1	5	Diversifying entrant; in vergelijking met 2006 veel van core activiteiten veranderd, omslagpunt ligt rondom 2012
1410	1	Ja;1/1	Onbevestigd	Onbevestigd	<20	2015	2015	2	0	2	4	

Web ID	Top25	TS-expertise	Innovatief	Startup-kenmerken	Werknemers	leeftijd o.b.v. web/LinkedIn	Wayback machine	TS-expertise	Innovatief	Startup	Validiteit-score	Type nieuwkomer
1487	1	Ja;1/1	"Ja;claimt innovatief te zijn;" "WESTRA onderscheidt zich doordat alles is gemaakt in Nederland; Van houten tafels en kasten tot de toeleveranciers en de metaalbewerkers." & " "Circulair ondernemen biedt positieve kansen; Het betekend een nieuwe manier van werken en daar werkt WESTRA 100% aan mee. De nadruk ligt op het voorkomen van afval en verspilling. Onze meubelfabriek is ingesteld op hergebruik van materialen en het zoveel mogelijk voorkomen van afvalstoffen. Zo worden onze houtspanen gebruikt voor het verwarmen van huizen en bedrijven, en houtafval wordt zo veel mogelijk in tafelbladen en poten gebruikt voor bepaalde modellen."	Nee; 2005 opgericht & "Zeven man heeft hij aan het werk. De helft geschoold meubelmaker, de andere helft Wajonger of Participatiewet; jongens met een zwerfverleden of een afkorting; Pdd-nos en datsoort dingen."	<20	2005	2009	2	2	0	4	
630	5	Ja;5/5	Ja; claimt innovatief te zijn; "enerGQ levert unieke en innovatieve oplossingen aan zowel de industrie, grote zakelijke instellingen als aan het MKB (midden- en kleinbedrijf). Voor deze oplossingen van energie besparen is ons bedrijf diverse keren beloond in het verleden (Innovatie Top 100, Accenture Innovation Award, IBM smart camp)."	Nee; 2006 opgericht en 3 medewerkers zichtbaar op LinkedIn	<20	2006	2011	2	2	0	4	
184	1	Ja;1/1	Ja; claimt innovatief te zijn; "Heel anders dan in de gangbare veehouderij, waar bij een aantal vleesrassen de koe zo uit haar krachten gegroeid is dat een kalf standaard met een keizersnede ter wereld moet komen. Onze manier van boeren staat dichterbij de natuur. De biologische veehouderij is een stap in de goede richting, maar gaat ons niet ver genoeg. Want zelfs biologisch rundvee wordt in de meeste gevallen toch in stallen afgemest, de laatste weken. Ook op gebied van soja gebruik (geen) en periode dat het kalf bij de koe loopt overstijgen we de normen, die gelden voor biologisch."	Ja; 2 medewerkers (en veel partners), opgericht 2012	<20	2012	2012	2	2	2	6	Geen Spinoff; "Toen we jaren geleden ontdekten wat voor bijzonder vlees de grazende natuurbeheerders opleverden, zijn we ons, naast het boer zijn, gaan richten op de verkoop van dit vlees."
581	1	Ja;1/1	Onbevestigd	Ja	<20	2013	2013	2	0	2	4	
68	1	Ja;0/1; +3 Creatieve Industrie + Energie + HTSM	Ja; claimt innovatief te zijn; "Wij zien een grote, bewegende markt voor ons. Een markt die continu vraagt om in te spelen op nieuwe ontwikkelingen. Denk aan combinaties met andere systemen in onze steeds meer verduurzamende maatschappij. En samenwerking met engineeringbureaus, lichtarchitecten en het er bij betrekken van de politiek. Bij gebruikers is veel kennis aanwezig en er leven steeds nieuwe ideeën op. In dit spanningsveld willen wij een rol spelen." & "Wij stellen onze doelstellingen regelmatig aan veranderende omstandigheden bij. Onze doelmarkten zijn scheepsbouw en metaal verwerkende industrie. Daarin wil Echtlicht eerst een toonaangevende rol gaan spelen in Noord-Nederland. Wij gaan ook een Nederlands distributiebedrijf oprichten en gaan op zoek naar andere landelijke partners. En wellicht in de toekomst ook verder, naar Duitsland, of België."	Ja; oprichting 2013 en 3 medewerkers zichtbaar op LinkedIn	<20	2013	2015	2	2	2	6	"Spinoff uit bedrijf; "Jan Erkelens is eigenaar van EchtLicht BV. Jan heeft managementfuncties gehad bij o.a. Heineken, Coca Cola, Smurfit en Abbott. Vanaf 2008 werkt hij als zelfstandig ondernemer. Kenmerkend voor hem is zijn passie voor productietechniek. Goed licht maakt productie makkelijker. Daarom ook een logische verbinding met EchtLicht. Maar er is ook een andere belangrijke factor die Jan succesvol maakt. En dat is een sterk ontwikkeld vermogen om problemen van technische aard, snel te kunnen doorgronden.""

Web ID	Top25	TS-expertise	Innovatief	Startup-kenmerken	Werknemers	leeftijd o.b.v. web/LinkedIn	Wayback machine	TS-expertise	Innovatief	Startup	Validiteit-score	Type nieuwkomer
382	2	Deels;2/2;Faciliterend/ Generiek	Onbevestigd	Twijfel;onderdeel van "blik op"	<20	2014	2014	1	0	1	2	
173	2	"Nee;0/2;False Positives opgegeven website blijkt niet meer online en ook niet terug te halen met waybackmachine, BENU Apotheek Pykstra Zwartemeer- Zwartemeer verwijst naar http://zwartemeer. benuapotheek.nl en blijkt grote keten te zijn"	Onbevestigd	Nee	nvt	nvt	niet gevon- den	0	0	0	0	
461	1	Ja;1/1	Onbevestigd	Nee;"Cars4all bestaat sinds oktober 2006"	nvt	2006	2011	2	0	0	2	
160	2	Ja;2/2	Onbevestigd	Onbevestigd	<20	2015	2015	2	0	2	4	
1155	2	Ja;2/2	Ja;"Wij zijn niet uw stand- daard bouwbedrijf Geen gebruikelijke bouwstenen maar keiharde program- meertaal en design. Public State bouwt innovatieve applicaties voor web en mobiel." &"Wij verdienen ons brood met innovatieve diensten en online producten." & "Wij zijn een jong bedrijf met geschoolde specialisten. Wij kennen de laatste trends en standaarden om het gebied van software. Hierdoor zijn wij in staat om unieke producten te ontwikkelen en continu te blijven innoveren."	Ja;oprichting 2012 en 4 medewerkers zichtbaar op LinkedIn	<20	2012	2016	2	2	2	6	Geen Spinoff; oprichter vanuit technische studie eerste bedrijven opgericht
1576	1	Ja;1/1	Onbevestigd	Ja;oprichting 2012 op basis van website	<20	2012	2015	2	0	2	4	
30	3	Deels;3/3;Uitvoering	Twijfel;maakt gebruik van innovatie maar lijkt zelf vooral een platform voor samenwerking en uitvoering;"Diary Campus Leeuwarden, een proefboerderij voor innovatie op het terrein van landbouw en veeteelt en onderdeel van de Uni- versiteit van Wageningen, heeft de mestvergisting nieuw leven in geblazen. Evergreen Energy heeft in samenwerking met de University of Minne- sota (VS) een innovatieve technologie ontwikkeld dat niet alleen zorgt dat de opbrengst van biogas uit biomassa hoger is, maar dat het ook cellulose / lignine volledig wordt vergist met betere kwaliteit biogas (meer methaan) tegen aanzienlijk lagere totale kosten."	Twijfel;Nieuwe samen- werking van verschillende initiatiefnemers, wel aan het einde van de in- novatieketen	<20		2016	1	1	1	3	
1051	2	Deels;2/2;Groothandel	Onbevestigd	Nee;Al sinds 1988 actief en inmiddels 7 vestigingen	nvt	1988	2001	1	0	0	1	
423	1	Ja;0/1;;+1 HTSM	Onbevestigd	Nee;oprichting 2007 en inmiddels 18 medewerkers	<20	2007	2013	2	0	0	2	
1121	3	Ja;3/3	Ja;claimt innovatief te zijn;"zijn wij uitgegroeid tot een unieke totaaldienst- verlener in utiliteitsbouw, inrichting en onderhoud van zakelijk onroerend goed met als specialisatie de medische sector"	Nee;"Sinds de oprichting in 2007"	<20	2007	2010	2	2	0	4	
1470	2	Ja;1/2;;;-1 Energie	Twijfel;gazellen award en ontwikkeld concepten en websiteontwerp.	Nee;opgericht in 2008 en 16 medewerkers	<20	2008	2008	2	1	0	3	
302	1	Nee;0/1;Groothandel	Onbevestigd	Onbevestigd	<20	2013	2013	0	0	2	2	

Web ID	Top25	TS-expertise	Innovatief	Startup-kenmerken	Werknemers	leeftijd o.b.v. web/LinkedIn	Wayback machine	TS-expertise	Innovatief Startup	Validiteit-score	Type nieuwkomer	
437	1	Ja;1/1;;+1 ICT	Twijfel;"Webdesign, techniek en online marketing. Alles onder één dak." & "Onze kracht ligt in het bedenken, maken en optimaliseren van digitale producten. Of het nu gaat om een eenvoudige 'one-pager' of een complex maatwerktraject, geef ons uw digitale probleem en wij bedenken de geschikte oplossing."	Ja;opgericht in 2015 en 5 medewerkers zichtbaar op LinkedIn	<20	2015	2016	2	1	2	5	Geen Spinoff; oprichter vanuit technische studie eerste bedrijven opgericht
1373	2	Ja;2/2	Ja;claimt innovatief te zijn;"he PoInt Game Controller is the latest revolution in Game Controllers. Not only did we make revolutionary changes to how a controller looks and feels, but we did a total, 180 degrees overhaul of any game controller out there! The patented PoInt Controller is unique in its features:"	Ja;oprichting 2015	<20	2015	2016	2	2	2	6	Geen Spinoff
1406	0	Ja;1/1	Twijfel;wetenschappelijke publicaties en boek uitgebracht	Nee;"Wij opereren sinds 1989 onder de naam Varenius"	<20	1989	2003	2	1	0	3	
117	3	Ja;2/3;;;-1 AgriFood	Ja;claimt innovatief te zijn;"Wij zijn een innovatieve onderneming gevestigd in Roden (Drenthe)"	Nee;2010 opgericht volgens LinkedIn	<20	2010	2012	2	2	0	4	"Na in de Food en Non-Food export gezeten te hebben is 15 jaar geleden de overstap gemaakt naar het ontwerpen en implementeren van software en het begeleiden van dergelijke trajecten. Onze achtergrond maakt Hunter-CRM zo flexibel."
1208	2	Deels;2/2;Faciliterend Generiek maar wel gespecialiseerd in fruitteelt, bloemteelt en melkveehouderij	Onbevestigd	Ja	<20	2012	2012	1	0	2	3	
1473	1	Ja;0/1;;+2 Creatieve Industrie + ICT	Twijfel;"CONCEPTS Bedenkers van creatieve en unieke projecten"	Ja;oprichting 2012 e 2 medewerkers zichtbaar op LinkedIn	<20	2012	2012	2	1	2	5	Geen Spinoff; LinkedIn-profiel van oprichter had voorheen 2 bedrijven in dezelfde sector
307	9	Deels;9/9;Faciliterend/Generiek, afgestemd op klanten in topsectoren	Onbevestigd	Onbevestigd	<20	2013	2013	1	0	2	3	
496	4	Ja;4/4	Ja;claimt innovatief te zijn;"COSMONIO's office is on the internet, since we always aim to work with the brightest minds in the world regardless of their location. Still, we need labs for our research and development activities, so we founded CORTEXLAB, HEALTHLAB and LIFELAB. Both of these labs are positioned in the heart of highly innovative campuses, where science, technology and creativity blend."	Ja;Team bestaat uit 11 medewerkers	<20	2013	2013; 2010 crawled, website had toen ander doeleinde	2	2	2	6	"Incumbent-backed entrant; ""COSMONIO is part of NVIDIA INCEPTION PROGRAM. Inception nurtures dedicated and exceptional startups who are revolutionizing industries with advances in AI and data science.""
808	3	Ja;3/3	Onbevestigd	Ja	<20	2013	2013	2	0	2	4	"2007- Softwarebedrijf MMSolutions geboren"(...)2013- Idee101 is geboren! Een nieuw concept met nieuwe visie" "In 2007 begonnen wij als MMSolutions. Vanuit zolderkamers en achtertuinen werkten wij toen onder andere aan websites en E-Commerce. Heel interessant, heel leuk, maar uiteindelijk niet waar wij het best in zijn. In 2010 daalde dat besef in en was de keus duidelijk: we gaan online software bouwen."

Web ID	Top25	TS-expertise	Innovatief	Startup-kenmerken	Werknemers	leeftijd o.b.v. web/LinkedIn	Wayback machine	TS-expertise	Innovatief	Startup	Validiteit-score	Type nieuwkomer
812	1	Ja;2/2	"Ja;claimt innovatief te zijn;" "Aan veel innovaties die in het werkveld hebben plaats gevonden heeft Ifmec een bijdrage geleverd. Ifmec heeft als kenniscentrum op het gebied van facility management in brede zin een eigen onderzoeksbureau en onderzoeksagenda. Al meer dan twintig jaar besteedt Ifmec aandacht aan onderzoek. De opgebouwde kennis wordt gebruikt om de opleidingen constant te blijven ontwikkelen."	Nee;oprichting 1994 en 30 medewerkers zichtbaar op LinkedIn	20+	1994	1998	2	2	0	4	
272	3	Ja;3/3	Ja;claimt innovatief te zijn;" "AIM Media ontwikkelde de laatste jaren verschillende applicaties die onze klanten het voeren van het bedrijf een stuk aangenamer, efficiënter of zelfs mogelijk hebben gemaakt. AIM Media ontwikkeld applicaties zowel in opdracht van klanten als in eigen beheer."&" "Dankzij onze creatieve geest lukt het bijna altijd de klant te ondersteunen met creatieve oplossingen, een concept en passende marketing."	Nee;oprichting 2008 en 13 medewerkers zichtbaar op LinkedIn	<20	2008	2013	2	2	0	4	;"AIM Media is opgericht in 2008 door Peter Bakker. In eerste instantie spreidde hij zijn creatieve geest tentoon naast zijn studie. Tegenwoordig doet hij dat Fulltime. Tjeuke en Peter zijn vanaf 2012 samen verder gegaan onder de vlag van AIM Media."
1572	3	Ja;3/3	Ja;claimt innovatief te zijn;" "wij zijn er om het online landschap elke dag een stukje mooier en slimmer te maken. Wij bedenken en ontwikkelen strategische design producten die zich onderscheiden van de massa."	Nee;opgericht in 2004 en 23 medewerkers	20+	2004	2004	2	2	0	4	
304	5	Ja;5/5	Ja;claimt innovatief te zijn;" "Arox staat voor innovatie. Met onze nieuwste innovaties zorgen we dat bedrijven klaar voor de toekomst zijn. De techniek van morgen, vindt u vandaag al bij ons. Neem een kijkje op de pagina van Arox Smart Technologies en leer onze autonome 'tel-drone' en de smartglasses kennen."	Nee;op basis van website al sinds 2008 actief	<20	te oud	2008; 2005 crawled maar geen info/foutmelding	2	2	0	4	
639	1	Nee;0/1;Groothandel	Onbevestigd	Nee;al minimaal sinds 2009 actief	nvt	te oud	2009	0	0	0	0	
883	4	Deels;4/4;Faciliterend/Generiek, afgestemd op klanten in topsectoren	Onbevestigd	Ja;Team bestaat uit 5 medewerkers	<20	2016	2016	1	0	2	3	
1322	1	Deels;1/1;Faciliterend/Generiek	Twijfel;expliciete vermelding: "Innovatief en altijd op zoek naar de best passende oplossing met de nieuwste innovatieve technieken." maar lijkt slechts gebruik te maken van innovaties zonder zelf een innovatieve bijdrage hieraan te leveren.	Ja;7 medewerkers volgens website	<20	2013	2017	1	1	2	4	
1542	1	Deels;1/1;Faciliterend/Generiek	Onbevestigd	Ja	<20	2013	2013	1	0	2	3	
304	5	Ja;5/5	Ja;claimt innovatief te zijn;" "Arox staat voor innovatie. Met onze nieuwste innovaties zorgen we dat bedrijven klaar voor de toekomst zijn. De techniek van morgen, vindt u vandaag al bij ons. Neem een kijkje op de pagina van Arox Smart Technologies en leer onze autonome 'tel-drone' en de smartglasses kennen."	Nee;op basis van website al sinds 2008 actief	<20	te oud	2008; 2005 crawled maar geen info/foutmelding	2	2	0	4	
370	2	Ja;1/2;;-1 Creatieve industrie	Onbevestigd	Nee	<20	te oud	2003	2	0	0	2	

Web ID	Top25	TS-expertise	Innovatief	Startup-kenmerken	Werknemers	leeftijd o.b.v. web/LinkedIn	Wayback machine	TS-expertise	Innovatief Startup	Validiteit-score	Type nieuwkomer	
1013	1	Ja;1/1	Ja;claimt innovatief te zijn;"MHB is innovatief en actief op het gebied van corporate events, live events, beursstands en videoschermen. Door de grote voorraad en de diverse disciplines is MHB de perfecte partner voor uw complete event."	Nee;"MHB Event Facilities heeft al ruim 10 jaar " Oprichting in 005 en 6 medewerkers zichtbaar op LinkedIn	<20	2005	2016	2	2	0	4	
1407	1	Ja;1/1	Ja;claimt innovatief te zijn;ontwikkeling van nieuwe games	Nee;oprichting 2002 met inmiddels 17 medewerkers zichtbaar op LinkedIn	<20	2002	2011; 2008 crawled, website had toen ander doeleinde	2	2	0	4	
1467	2	Ja;1/2;;;-1 Energie	Twijfel;"Warbout Distributie is geen 13 in een dozijn distributeur. Onze keuze voor merken en producten is weloordacht en onderscheidend. (...) De markt, vraag naar producten en de techniek verandert snel. Warbout Distributie heeft een enthousiast team mensen dat telkens probeert in te spelen op deze veranderingen. Zo zijn wij al ruim 2 jaar druk bezig met het verkoop van softwarelicenties via ons portal, ontwikkelen we onze e-commerce mogelijkheden zelf en waar het producten betreft bemoeien we ons gevraagd en ongevraagd zelfs met de productie!"	Nee;LinkedIn data wijst op een voormalig medewerker die sinds 2001 werkzaam zou zijn bij Warbout	<20	te oud	2015	2	1	0	3	
596	4	Ja;4/4	Ja;claimt innovatief te zijn;"Our range of products varies from NP reducing Biopellets, Easy feed, Reef pearls, Joes Juice, Coral Protect and Ocean Prime Copepods, which we develop and manufacture in the Netherlands."	Nee;Oprichting 2001 volgens LinkedIn profiel oprichter	nvt	te oud	2004	2	2	0	4	
1095	5	Ja;4/5;;;-1 AgriFood	Ja;claimt innovatief te zijn;"JONG, INNOVATIEF EN BETROUWBAAR"	Ja;2015 opgericht en 2 medewerkers zichtbaar op LinkedIn	<20	2017	2017	2	2	2	6	Diversifying entrant; LinkedInprofielen van de oprichters met werkervaringen in enkel de marketing en ICT.
956	1	Deels;1/1;Producent Sanitair meer gericht op design dan op "Duurzaam functioneren van watersystemen";+1 Creatieve Industrie	Ja;claimt innovatief te zijn;"Het innovatieve sanitairmerk LooX kenmerkt zich door..."	Nee;locatie Deventer	nvt	nvt	2018; 2005 crawled maar geen info/ foutmelding nu redirect naar https://loox.nl/nl/	1	2	0	3	
1129	2	Deels;2/2;Faciliterend/Generiek	Onbevestigd;vooral dealer en advies	Ja	<20	2011	2011	1	0	2	3	
362	1	Ja;1/1	Twijfel;"BeterBrein is een organisatie gericht op het ontdekken en verspreiden van innovaties in de geestelijke gezondheidszorg."	Nee;linkedin eigenaar al sinds 2007	<20	2007	2008	2	1	0	3	
578	1	Nee;0/1; Nieuwsplatform	Onbevestigd	nvt	nvt	nvt	2011	0	0	0	0	
598	1	Ja;1/1;+2 Creatieve Industrie + ICT	Onbevestigd	Nee;opgericht in 2000 en 3 medewerkers zichtbaar op LinkedIn	<20	2000	2018	2	0	0	2	
883	4	Deels;4/4;Faciliterend/Generiek, afgestemd op klanten in topsectoren	Onbevestigd	Ja;Team bestaat uit 5 medewerkers	<20	2016	2016	1	0	2	3	
1319	1	Deels;1/1;Faciliterend/Generiek	Onbevestigd	Nee;65 medewerkers volgens website en maar liefst 115 zichtbaar op LinkedIn. Oprichting 2005 volgens linkedin	20+	2005	2006	1	0	0	1	
1570	2	Ja;2/2;;+1 ICT	Twijfel;Business intelligence en webdevelopment	Nee;"Met een team van 25 collega's " en volgens linkedin oprichting in 2008 en inmiddels 54 werknemers zichtbaar op LinkedIn	25+	2008	2008	2	1	0	3	
1188	1	Ja;1/1	Onbevestigd	Nee;niet meer levensvatbaar;"Wegens Faillissement is dit bedrijf gesloten."	<20	2013	2013	2	0	0	2	

Web ID	Top25	TS-expertise	Innovatief	Startup-kenmerken	Werknemers	leeftijd o.b.v. web/LinkedIn	Wayback machine	TS-expertise Innovatief Startup	Validiteit-score	Type nieuwkomer		
1233	5	Ja;5/5	Ja;claimt innovatief te zijn;"We develop and manufacture high-tech measurement instruments for field operations&"The Sensor Factory is at home in all stages of sensor technology development, from concept to field-testing to implementation and roll-out."	Ja;2015 opgericht en 2 medewerkers zichtbaar op LinkedIn	<20	2015	2017	2	2	2	6	Spinoff uit bedrijf; De oprichters hebben volgens het LinkedIn profiel sector gerelateerde ervaring in de R&D tak van gevestigde bedrijven."The Sensor Factory first came together over a decade ago, searching for new ideas in untapped sensor technologies and undiscovered applications.(...) The years following saw several of these bright ideas taken to the next level, becoming prototypes and test cases. And finally, sales to multinational companies." "The Sensor Factory is made up of a team of diverse, innovative professionals intent on bringing the laboratory to the process.(...) Growing and expanding into the future."
1310	2	Ja;2/2	Ja;claimt innovatief te zijn;"Taal&Digitaal staat voor creatieve internet-technologie. Daarbij gaan we grote uitdagingen niet uit de weg. Alles met het doel om de digitale wereld te verrijken"	Twijfel;zelfde eigenaar al in 2006 maar andere activiteiten, mogelijk sinds 2012 diversifying entrant, inmiddels 8 werknemers	<20	2006	2006	2	2	1	5	Diversifying entrant; in vergelijking tot 2006 veel van core activiteiten veranderd, omslagpunt ligt rondom 2012
1374	3	Ja;1/3;software oplossingen voor computersystemen maar geen ontwikkeling van systemen en materialen zelf;;-1 HTSM	Ja;claimt innovatief te zijn;ontwikkelt en onderhoudt nieuwe producten	Ja;oprichting 2012 met 1 medewerker zichtbaar op LinkedIn, aangezien het om een co-owner gaat zou er ook nog minimaal een 2e werknemer moeten zijn waarvan geen LinkedIn profiel zichtbaar is	<20	2012	2012	2	2	2	6	Spinoff uit onderzoek-instelling; "As part of the Distributed Systems research group I assisted in the research, design, development and evaluation of office energy saving through component based automation." (linkedin)
1392	4	Deels;4/4;Faciliterend/Generiek, delen en bundelen van online informatie uit topsectoren	Onbevestigd	Onbevestigd	<20	2013	2013	1	0	2	3	
38	2	Ja;2/2;;+1 LS&H	Ja;claimt innovatief te zijn;"In 2013 ontwikkelden wij hiervoor de eerste ademhalingscoach applicatie die wereldwijd meer dan 32.000 keer is gedownload. In 2017 zijn wij gestart met de ontwikkeling van een persoonlijk coaching- en feedback systeem bestaande uit een sensor die meet hoe je buikademhaling is, de frequentie en diepte van je ademhaling en een applicatie voor een mobile device voor zien van een oefenpakket om je ademhaling te verbeteren of te herstellen."	Ja;2015 opgericht en 4 medewerkers zichtbaar op LinkedIn en 6 medewerkers werkzaam volgens de website	<20	2015	2017	2	2	2	6	Diversifying entrant; oprichters voorheen Stress counselor/ademhalingstherapeuten: "Wij, Nellie Bontekoe en Sonja Nijhuis, beiden ervaringsdeskundigen en geschoolde counselors, hebben in onze praktijk voor ademtherapie en counseling, ruim 14 jaar mensen geholpen die te maken hadden met een disfunctionele ademhaling, vaak ontstaan door stress, paniek of angst. Graag wilden wij meer mensen met een disfunctionele ademhaling bereiken en ze met de juiste gereedschappen in staat stellen, hun ademhaling te verbeteren."
604	4	Deels;9/9;Faciliterend/Generiek	Onbevestigd;economisch samenwerkingsverband	Nee; Samenwerkingsverband	nvt	nvt	2013	1	0	0	1	
1136	2	Nee;0/2;Groothandel	Onbevestigd	Nee;al minstens sinds 2003 actief	nvt	te oud	2003	0	0	0	0	
1244	3	Deels;3/3;Faciliterend/Generiek	Onbevestigd	Ja;Geen LinkedIn gevonden	<20	2012	2012	1	0	2	3	
200	1	Ja;0/1;;+1 Creatieve Industrie;-1 Energie	Onbevestigd	Ja	<20	2013	2013	2	0	2	4	

Web ID	Top25	TS-expertise	Innovatief	Startup-kenmerken	Werknemers	leeftijd o.b.v. web/LinkedIn	Wayback machine	TS-expertise	Innovatieve Startup	Validiteit-score	Type nieuwkomer	
214	3	Nee;0/3;Domein buitengebruik of gebruikt voor andere doeleinde	Onbevestigd	Nee;Website buiten gebruik	nvt	nvt	2001; sinds 2001 crawled, verschillende doeleinde domeinnaam regelmatig buiten gebruik	0	0	0	0	
239	2	Ja;2/2	Onbevestigd	Nee;"De Van der Woude Groep is in de jaren tachtig begonnen met agrarisch loonwerk en egalisatie. In het begin van de negentiger jaren werd uitgebreid met de verhuur van grondverzetmachines, te weten een hydraulische graafmachine van 20 ton en een bulldozer. Inmiddels is dit machinepark uitgebreid, variërend van graafmachines van 5 tot 125 ton, dumpers, bulldozers en laadschoppen. Begin 2000 zijn de werkzaamheden binnen de Van der Woude Groep uitgebreid met baggerwerkzaamheden."	nvt	te oud	2011	2	0	0	2	
1525	2	Deels;1/2;Vooral vastgoed ontwikkeling soms gericht om duurzaamheid;;-1 LS&H	Ja;claimt innovatief te zijn;"Waarborg Vastgoed is huisvester van organisaties die het verschil maken in hun branche of in de maatschappij. Een vastgoed-partner van succesvolle ondernemers die verder kijken, anders denken en innovatieve ideeën werkelijkheid maken." & "To invest, house and facilitate in a sustainable way. That is, in a nutshell, the ambition of Waarborg Vastgoed. (...) Every day, our team is committed to ensure we provide sustainable, inspiring and innovative work spaces for our residents."	Nee;Linkedin mede eigenaar al sinds 2007 in dienst, inmiddels 19 medewerkers zichtbaar op LinkedIn	<20	te oud	2012	1	2	0	3	
277	1	Ja;1/1;;+1 ICT	Ja;claimt innovatief te zijn;"Ondernemer in digitale innovatie (...) heb ik de Digitectuur visie ontwikkeld. Voor digitale innovatie in de fysieke ruimte. Om fysiek en digitaal te verbinden. Samen met mijn vrouw Selma (cultuur socioloog) ben ik Platform Digitectuur.nl gestart."	Ja;oorspronkelijke bedrijf bestaat al sinds 2010 maar sinds 2015 "oprichter / directeur at netwerk Digitectuur.nl" (linkedin) inmiddels 3 medewerkers zichtbaar op LinkedIn	<20	2015	2002; website eerder gebruikt voor andere doeleinde	2	2	2	6	Geen Spinoff; oprichter voorheen al eerder begonnen met bedrijf in dezelfde sector
464	3	Ja;3/3	Ja;claimt innovatief te zijn;Ontwikkelt eigen game	Ja	<20	2015	2015	2	2	2	6	Geen Spinoff;Althans op zowel website als LinkedIn is hierover niks te vinden.
518	1	Ja;2/2	Ja;claimt innovatief te zijn;"DBK gelooft in het maximale halen uit online business. DBK wil opdrachtgevers laten groeien! Wij geloven ook in snelle en persuasieve websites en verder denken dan de rest. Daarom legt DBK voor opdrachtgevers de essentiële basis voor een succesvolle toekomst." & "Wat is er zo uniek aan ons? Wij sturen op basis van cijfers, en niet op basis van gevoel of ervaringen. We zorgen ervoor dat alle online resultaten meetbaar zijn en maken beslissingen op basis van data uit het verleden. Op die manier zorgen we voor meer bezoekers, verhogen we de orderwaarde én conversies. Omdat we alles meetbaar maken, kunnen opdrachtgevers precies zien hoe campagnes hebben gepresteerd. Wel zo prettig. (...) Wij maken (en houden) onze opdrachtgevers graag tevreden. Wij zijn partners van onze opdrachtgever en ze betalen ons pas als ZIJ succes hebben. "	Ja;oprichting 2012 en inmiddels 9 medewerkers	<20	2012	2013	2	2	2	6	Spinoff uit bedrijf; Topondernemer en oprichter heeft 21 bedrijven (blijkt uit interview op linkedin, bedrijven zelf staan lang niet allemaal onder zijn naam in linkedin) en is al decennia lang ondernemer, werkervaringen bij Google Amsterdam worden benoemd als cruciaal voordeel voor succesvol ondernemen.

Web ID	Top25	TS-expertise	Innovatief	Startup-kenmerken	Werknemers	leeftijd o.b.v. web/LinkedIn	Wayback machine	TS-expertise	Innovatief Startup	Validiteit-score	Type nieuwkomer
612	1	Nee;0/1;klusbedrijf	Onbevestigd	Nee;"Gespecialiseerd in alle voorkomende tegelwerken voor particulieren en bedrijven met meer dan 35 jaar ervaring in deze branche."	nvt	te oud	2003	0	0	0	0
654	4	Ja;3/4;;-1 ICT	Ja;claimt innovatief te zijn;Innovatieaward winner	Nee;2005 opgericht, 18 medewerkers	<20	2005	2008	2	2	0	4
846	3	Ja;3/3	Ja;claimt innovatief te zijn;"Innovative Technical solutions" & "Vanuit Drachten ontwikkelt ITS innovatieve en technologisch geavanceerde oplossingen voor bedrijven die zelf onvoldoende ontwikkelcapaciteit of kennis in huis hebben."	Ja;Linkedin:" Year founded 2012 (...) Company size 2-10 employees"	<20	2012	2012	2	2	2	6
1222	3	Ja;3/3	Ja;claimt innovatief te zijn;"Onze ontwerpen zijn energiezuinig, innovatief...."	Nee;"De SchreuderGroep is adviesbureau met 50 jaar ervaring." Linkeding oprichtings jaar 2004	20+	2004	1998	2	2	0	4
1392	4	Deels;4/4;Faciliterend/Generiek, delen en bundelen van online informatie uit topsectoren	Onbevestigd	Onbevestigd	<20	2013	2013	1	0	2	3
786	2	Ja;2/2	Ja;claimt innovatief te zijn;Benoemt niet heel expliciet de innovatie maar wel de nieuwe producten en de wetenschappelijke expertise die ingezet wordt voor de ontwikkeling en productie: "New! 2d transfer system. Designed and assembled at HQ Graphene. (...) This manual transfer system is designed to accurately place a stamp on a substrate, which can be moved relative to each other not only in the x-, y- and z-directions, but also by tilting and rotating. This allows the highest degree of freedom in the alignment of crystal flakes when fabricating hetero-structures. The substrate is fixated with vacuum and can be heated during the process. The compact design allows for easy placement inside a glove box, providing the option to work with air sensitive materials." & "Our scientific staff has the expertise in the production and characterization of large size and high purity single crystals. The crystals are characterized with X-ray diffraction (single crystal XRD and powder XRD), Energy-dispersive X-ray spectroscopy (EDX), Raman and additionally with Hall measurements. We can grow all kinds of two dimensional crystals."	Ja;oprichting 2012	<20	2012	2013	2	2	2	6
15	1	Ja;1/1;;+1 HTSM	Ja;claimt innovatief te zijn;"Met innovatieve drones verzamelt Agrifly informatie over het gewas. Deze drones zijn voorzien van geavanceerde sensoren die informatie vastleggen die het menselijk oog niet kan waarnemen. Multispectrale beelden worden in vier golf lengtes vastgelegd die speciaal zijn geselecteerd om de reflectie van het gewas te meten te weten groen, rood, nabij infrarood (NIR) en nabij rood (Red Edge)."	Ja;oprichting 2016 en 2 medewerkers zichtbaar op LinkedIn	<20	2016	2017	2	2	2	6
											Spinoff uit bedrijf;"Patrick founded Fairtrasa in 2005 after a personal epiphany led him to devote his life to social entrepreneurship." (...) "He started his business career with the global consultancy firm Deloitte and was hired in 1999 by one of his clients, Glencore, the largest commodities trading firm in the world, where he quickly established himself as a Department Head." "Since 2005, Fairtrasa has pioneered a unique and innovative solution to these global challenges, impacting the lives of over 40,000 people, including farmers and their families. Our objective now is to scale this solution to reach millions more farmers worldwide."
											Spinoff uit bedrijf;"With 22 years experience in defense (marines and airforce) I am an initiative and assertive person, a "quick thinker", able to analyse company operations, tactics and strategies because of my broad experience at all company levels. "
											Spinoff uit onderzoeksinstelling; Volgens LinkedIn: PhD in Applied Physics: "As a PhD student I performed spin transport measurements in graphene and carbon nanotubes. & R&D Experimental Physicist bij MAPPER lithography (gevestigd bedrijf 200+FTE in Lithografie, Chipmanufacturing, E-beam-technologie en optica) en als postdoc nog steeds verbonden en onderzoeksinstellingen.
											Diversifying entrant; Jonge student werktuigbouwkunde en technische bedrijfskunde met voorheen werkervaring in de akkerbouwen en grondrijden en inmiddels letterlijk op de zolderkamer begonnen met ondernemen.

Web ID	Top25	TS-expertise	Innovatief	Startup-kenmerken	Werknemers	leeftijd o.b.v. web/LinkedIn	Wayback machine	TS-expertise	Innovatief Startup	Validiteit-score	Type nieuwkomer	
27	1	Ja;1/1;;+1 Logistiek	Onbevestigd	Nee;"Transportbedrijf Betten is in 1930 opgericht door grootvader Jannes Betten. "	nvt	1930	2004	2	0	0	2	
370	2	Ja;1/2;;;-1 Creatieve industrie	Onbevestigd	Nee	<20	te oud	2003	2	0	0	2	
1556	1	Ja;1/1	Twijfel;wel onderscheidend door duurzaamheid en verantwoord:"Dag op dag zetten wij ons met onze natuurinclusieve boerderij in om de wereld te helpen een stukje beter en duurzamer te maken." & "Door het vlees van de Groninger Blaarkop te consumeren, draag je juist bij aan de instandhouding van het prachtige zeldzame koeienras."	Ja;oprichting vlgens LinkedIn oprichting in 2015 en 2 medewerkers zichtbaar op LinkedIn	<20	2015	2017	2	1	2	5	Geen Spinoff; Eigenaar al eerder eigenaar (vermoedelijk overgenomen) bij bedrijven in andere sector (bewerken en vervaardigen van elektronische producten)
729	2	Deels;2/2;Faciliterend Generiek	Onbevestigd;Wel onderscheidend:"HABITAT advocatenkantoor is het enige kantoor in Nederland dat louter is gespecialiseerd in omgevingsrecht."	Nee;"We zijn al meer dan 10 jaar actief" en nu 5 medewerkers	<20	te oud	2010	1	0	0	1	
811	8	Nee;0/8;Bemiddeling	Onbevestigd	Onbevestigd	<20	2013	2013; inmiddels offline	0	0	2	2	
507	2	Deels;0/2;Groothandel; +1 Logistiek	Onbevestigd	Nee;Groothandel in heel Noord Europa	nvt	nvt	2017	1	0	0	1	
565	1	Ja;1/1	Onbevestigd	Nee;"meer dan 14 jaar eraring" en 9 medewerkers volgen websiteinformatie	<20	te oud	2010	2	0	0	2	
1051	2	Deels;2/2;Groothandel	Onbevestigd	Nee;Al sinds 1988 actief en inmiddels 7 vestigingen	nvt	1988	2001	1	0	0	1	
1450	1	Deels;1/1;vooral financiële diensten	Onbevestigd	Nee;"Volkswagen Financial Services is een handelsnaam van Volkswagen Pon Financial Services B.V., DFM N.V., DFM Verzekeringen B.V. en Volkswagen Leasing B.V. "	nvt	te oud	2007	1	0	0	1	
453	1	Ja;1/1;;+1 Creatieve Industrie	Ja;claimt innovatief te zijn;"een hoog kennisniveau en constante innovatie zijn de succesfactoren van het Avepro-team." & "De A/V-technologie staat nooit stil. Daarom innoveert Avepro continu. Sterker nog: innovatie is onze missie."	Ja;oprichting 2016 en 2 medewerkers zichtbaar op LinkedIn	<20	2016	2017	2	2	2	6	Diversifying entrant;Combinatie van twee verschillende achtergronden:"Avepro ontstond in 2016 vanuit STEALTH elektrotechniek (Suward de Jong) en SONAR6 licht, geluid en video (Bert Zonderland)."
1123	1	Ja;1/1	Ja;claimt innovatief te zijn;"Want naast codetaal houden we ook van innoveren."	Nee;"Daarbij komt PHP Solutions van pas, dat al bijna vijftien jaar bestaat."	<20	te oud	2008	2	2	0	4	
1524	1	Ja;1/1;;+1 Creatieve Industrie	Ja;claimt innovatief te zijn;"Kernwaarden (...) Innovatief"	Nee;oprichting 2001 en 9 medewerkers zichtbaar op LinkedIn	<20	2001	2002	2	2	0	4	
830	3	Ja;3/3	Ja;claimt innovatief te zijn;Focus op nieuwe markten en onderscheiden: "As a result of the country's intense focus on mass production, China's economy grew dramatically during the early 2000s and a burgeoning market for luxury items such as aftermarket wheels developed quickly alongside it. With newfound disposable income at their fingertips, China's rapidly emerging middle class propelled sales of INOVIT products to a level equal to that in Europe within a mere two months." & "The fifth INOVIT operation was established in Northern Italy in 2010 (...) Italy is home to some of the biggest names the alloy wheel world has ever known and while competition was tough, we were determined to set ourselves apart from the crowd with our unique designs and expert engineering."	Nee	<20	1998	2004; 2000 crawled, maar geen info	2	2	0	4	"January 13 1998 (...) The Start (...) Established by a sole trader early in 1998, the business commenced operation from a home office and small warehouse located in Friesland, a northern province in the Netherlands. After successfully sourcing suppliers for wheels and tires, the business began trading and a customer network that stretched across Northern Europe and the UK started to develop. Tenacity and a passion for business quickly led to a sharp rise in interest in the company's products and services and it soon became apparent that expansion was the necessary next step. With a growing customer base and great reputation, it was time to source a much bigger warehouse and take on some extra staff to cope with demand."

Web ID	Top25	TS-expertise	Innovatief	Startup-kenmerken	Werknemers	leeftijd o.b.v. web/LinkedIn	Wayback machine	TS-expertise	Innovatief	Startup	Validiteit-score	Type nieuwkomer
1335	1	Ja;1/1;;+2 AgriFood + Chemie	Ja;claimt innovatief te zijn;"Daarvoor hebben wij een kenniscentrum opgericht. Wij hebben inmiddels een grote database opgebouwd met resultaten van testen op allerlei materiaalsoorten. Ons kenniscentrum stelt ons in staat u een juist advies geven. Op deze pagina ziet u een greep uit onze projecten en testen."	Onbestigd;"The Coating Company heeft reeds meer dan 5 jaar ervaring in de food-sector."	<20	2017	2017	2	2	2	6	Onbekend
239	2	Ja;2/2	Onbevestigd	"Nee;"De Van der Woude Groep is in de jaren tachtig begonnen met agrarisch loonwerk en egalisatie. In het begin van de negentiger jaren werd uitgebreid met de verhuur van grondverzetmachines, te weten een hydraulische graafmachine van 20 ton en een bulldozer. Inmiddels is dit machinepark uitgebreid, variërend van graafmachines van 5 tot 125 ton, dumpers, bulldozers en laadschoppen. Begin 2000 zijn de werkzaamheden binnen de Van der Woude Groep uitgebreid met baggerwerkzaamheden."	nvt	te oud	2011	2	0	0	2	
631	1	Twijfel;1/1;"We repair worldwide marine and industrial diesel engines": onduidelijk of dit volgens de topsector energie wel of niet bijdraagt voor de transitie naar een betaalbaar, betrouwbaar en duurzaam energiesysteem.	Onbevestigd	Nee;oprichting 2008 en 3 medewerkers op linkedin (bedrijfsinformatie geeft 10-50 medewerkers totaal aan)	<20	2008	2009	1	0	0	1	
1293	1	Nee;0/1	Onbevestigd	Nee;al zekers sinds 2001 actief	nvt	te oud	2001	0	0	0	0	
522	3	Deels;3/3;Uitvoering	Onbevestigd	Nee;al sinds 2002 actief volgens website	nvt	2002	2002	1	0	0	1	
644	2	Ja;2/2	Twijfel;"CONCEPTS Bedenkers van creatieve en unieke projecten"	Ja;2014 opgericht volgens linkedin en 4 medewerkers zichtbaar op LinkedIn. Volgens website: "Het bureau telt in totaal 6 vaste medewerkers en maakt daarnaast op projectmatige basis gebruik van een aantal vertrouwde ontwerp bureaus."	<20	2014	2013;2000 crawled, website had toen ander doeleinde	2	1	2	5	Spinoff uit bedrijf;oprichters voorheen werkzaam bij gevestigde bedrijven in aanverwante sector
848	1	Ja;1/1;;+1 ICT	Ja;claimt innovatief te zijn;ontwikkelt nieuwe applicaties voor mobiel en web	Onbevestigd;zowel op de website als linkedin geen informatie over medewerkers en oprichting	<20	2016	2016	2	2	2	6	Onbekend
1227	2	Ja;1/1;;+1 ICT	Ja;claimt innovatief te zijn;ontwikkelt eigen "ScreenFeed producten" voor verschillende doeleinden: "Waar ScreenFeed Dynamic de adverteerder ondersteunt, voorziet Information uw bezoekers in één oogopslag van allerlei belangrijke tekstuele informatie.	Nee;"ScreenFeed is opgericht in 2005." en tevens onderdeel van het al langer bestaande slash2media	nvt	2005	2007;inmiddels offline	2	2	0	4	
859	4	Ja;4/4;Groothandel met revisie service	Onbevestigd	Nee;2007 al operationeel (waybackmachine)	nvt	nvt	2007	2	0	0	2	
669	2	Ja;1/2;;+1 Creatieve Industrie;-1 AgriFood	Ja;claimt innovatief te zijn;"Nu, een paar jaar verder, is dit project uitgegroeid tot het grootste online sport- en gezondheidsplatform van Nederland en België. We onderscheiden ons van andere websites door de redactie die bestaat uit louter professionals, zoals wetenschappers, sportdietisten en fysiotherapeuten." en ontwikkelt eigen producten en diensten	Ja;2012 opgericht en 22 medewerkers zichtbaar op LinkedIn- vermoedelijk vooral parttime gezien de combinatie met andere (fulltime) banen	20-25	2012	2012	2	2	2	6	Geen Spinoff;"We begonnen FIT.nl vier jaar geleden, als een paar studenten die zich stoorden aan de hoeveelheid onjuiste informatie over sport en gezondheid op het internet. Als jonge enthousiastelingen willen we de vele sportfabeltjes de wereld uit helpen en daarom richtten we een website op met gratis wetenschappelijk onderbouwde tips."

Web ID	Top25	TS-expertise	Innovatief	Startup-kenmerken	Werknemers	leeftijd o.b.v. web/LinkedIn	Wayback machine	TS-expertise	Innovatief Startup	Validiteit-score	Type nieuwkomer	
1104	3	Nee;0/3;Bemiddeling	Onbevestigd	Onbevestigd	<20	2016	2016	0	0	2	2	
307	9	Deels;9/9;Faciliterend/ Generiek, afgestemd op klanten in topsectoren	Onbevestigd	Onbevestigd	<20	2013	2013	1	0	2	3	
1015	5	Ja;5/5	Ja;claimt innovatief te zijn;"We offer (...) Development of innovative solutions"	Onbevestigd	<20	2013	2013	2	2	2	6	Incumbent-backed entrant; "Microbial Analysis is part of the company Bioclear bv. Bioclear was established in 1988 (...) That is why in 2001 we started to offer our microbial analyses commercially to third parties (...) The number of analyses steadily increased and in 2012 (...) The formation of Microbial Analysis by Bioclear was a fact."
805	1	Ja;1/1	Ja;claimt innovatief te zijn;"In 2015 zijn wij ook Fleetprint gestart om oplossingen voor de mobiliteitsmarkt te ontwikkelen zoals auto's die u op afstand kunt ontgrendelen. Een ideale oplossing voor poolauto's. Kortom ICEU/ Fleetprint is een innovatief bedrijf wat altijd open staat voor nieuwe en betere oplossingen om het voor de berijder of uw wagenpark zo prettig mogelijke te maken."	Ja;opgericht in 2015 en 2 medewerkers zichtbaar op linkedin (bedrijfsinformatie geeft 11-50 medewerkers totaal aan). "In 2006 is Inbouw Centrum Europa, ICEU in het kort, in Groningen opgericht" maar "In 2015 zijn wij ook Fleetprint gestart om oplossingen voor de mobiliteitsmarkt te ontwikkelen zoals auto's die u op afstand kunt ontgrendelen. Een ideale oplossing voor poolauto's. Kortom ICEU/Fleetprint is een innovatief bedrijf wat altijd open staat voor nieuwe en betere oplossingen om het voor de berijder of uw wagenpark zo prettig mogelijke te maken."	<20	2015	2009	2	2	2	6	Spinoff uit bedrijf;oprichters hebben werkzaamheden uitgebreid om ook nieuwe producten te ontwikkelen
246	3	Ja;3/3	Ja;claimt innovatief te zijn;"Bij 8D Games wordt veel gebruik van brainstormsessies om ideeën te bedenken (...) om creatieve oplossingen te verzinnen."	Ja;oprichting 2012 en 15 medewerkers zichtbaar op LinkedIn	<20	2012	2013	2	2	2	6	Diversifying entrant; Oprichter volgens linkedin voorheen met werkervaring in andere sector(en)
421	2	Nee;0/2;Bemiddeling	Onbevestigd	Nee;Oprichting 2010 en 17 medewerkers zichtbaar op LinkedIn	<20	2010	2011	0	0	0	0	
840	4	Ja;3/4;;;-1 Water	Ja;claimt innovatief te zijn;"In 2012 we decided to focus our Research & Development and portfolio extension on Women's Health. This resulted in several patents and patent applications in the field of prognostic assays to determine the risk and pregnancy related disorders."&"Unique and ready-to-use antibodies and complete innovative diagnostic kits developed for all currently available flow cytometers."	Nee;opgericht in 1987	nvt	te oud	1999	2	2	0	4	;"It was founded in cooperation with the leading University of Groningen, in the north of The Netherlands in 1987."
1233	5	Ja;5/5	Ja;claimt innovatief te zijn;"We develop and manufacture high-tech measurement instruments for field operations&"The Sensor Factory is at home in all stages of sensor technology development, from concept to field-testing to implementation and roll-out."	Ja;2015 opgericht en 2 medewerkers zichtbaar op LinkedIn	<20	2015	2017	2	2	2	6	Spinoff uit bedrijf; De oprichters hebben volgens het LinkedIn profiel sector gerelateerde ervaring in de R&D tak van gevestigde bedrijven."The Sensor Factory first came together over a decade ago, searching for new ideas in untapped sensor technologies and undiscovered applications.(...) The years following saw several of these bright ideas taken to the next level, becoming prototypes and test cases. And finally, sales to multinational companies." "The Sensor Factory is made up of a team of diverse, innovative professionals intent on bringing the laboratory to the process.(...) Growing and expanding into the future."

Web ID	Top25	TS-expertise	Innovatief	Startup-kenmerken	Werknemers	leeftijd o.b.v. web/LinkedIn	Wayback machine	TS-expertise	Innovatief	Startup	Validiteit-score	Type nieuwkomer
604	4	Deels;9/9;Faciliterend/Generiek	Onbevestigd;economisch samenwerkingsverband	Nee; Samenwerkingsverband	nvt	nvt	2013	1	0	0	1	
1244	3	Deels;3/3;Faciliterend/Generiek	Onbevestigd	Ja;Geen LinkedIn gevonden	<20	2012	2012	1	0	2	3	
1268	2	Nee;0/2;Groothandel	Onbevestigd	Nee;"Met ruim 25 jaar ervaring in het vak bent u bij ons aan het juiste adres voor een duurzaam zonnepaneelensysteem."	nvt	te oud	2012	0	0	0	0	
116	3	Ja;4/4	Ja;claimt innovatief te zijn;Focus ligt vooral op toepassing van 3D print technologie voor Productontwerp Product ontwikkeling en bemiddeling	Ja;Team bestaat uit 2 medewerkers	<20	2015	2015	2	2	2	6	Spinoff uit bedrijf; Oprichters volgens LinkedIn voorheen werkzaam in dezelfde sector. Website:"Hulotech 3D printing & Engineering is een jong en flexibel bedrijf welke gerund wordt door 2 eigenaren welke beide een brede technische achtergrond hebben in ontwikkeling, productie en kwaliteits-management. We hebben ruim 15 jaar ervaring in het ontwerpen van complexe 3D modellen middels geavanceerde 3D CAD software. Door de jarenlange kennis en ervaring binnen de High-Tech maakindustrie (Medische-, defensie- en industriële sector)..."
171	2	Ja;2/2	Ja;claimt innovatief te zijn;"het complete proces van idee tot model. In nauwe samenwerking (...) en onder strikte geheimhouding." &"Ontwerp- en engineeringbureaus, productiebedrijven en R&D afdelingen van gerenommeerde bedrijven uit verschillende branches zijn al jaren onze klant. Wij denken mee en adviseren bij de ontwikkeling van producten gebaseerd op onze brede ervaring. Zo werkt Modelmakerij Promoac samen met de klant tijdens het ontwikkelingsproces." &"Wij zijn altijd bezig onze processen te verbeteren of nieuwe mogelijkheden toe te passen om dat hoge niveau te handhaven."	Nee;"Sinds 1996 verzorgt Modelmakerij Promoac het complete proces van idee tot model. "	nvt	1996	2005	2	2	0	4	
496	4	Ja;4/4	Ja;claimt innovatief te zijn;"COSMONIO's office is on the internet, since we always aim to work with the brightest minds in the world regardless of their location. Still, we need labs for our research and development activities, so we founded CORTEXLAB, HEALTHLAB and LIFELAB. Both of these labs are positioned in the heart of highly innovative campuses, where science, technology and creativity blend."	Ja;Team bestaat uit 11 medewerkers	<20	2013	2013; 2010 crawled, website had toen ander doeleinde	2	2	2	6	"Incumbent-backed entrant; ""COSMONIO is part of NVIDIA INCEPTION PROGRAM. Inception nurtures dedicated and exceptional startups who are revolutionizing industries with advances in AI and data science."" "
549	1	Ja;1/1;;+1 ICT	Ja;claimt innovatief te zijn;"Het internet moet mooier! Dit is de missie van de Staat van Creatie. Ons middel is creatief conceptdenken en ultiem webdesign. Dat betekent een creatief, onderscheidend, uniek, conceptmatig, gebruiksvriendelijk, responsive en doeltreffend webdesign."	Ja;opgericht 2011 en 12 medewerkers zichtbaar op LinkedIn en 13 medewerkers volgens de website.	<20	2011	2011	2	2	2	6	Spinoff uit bedrijf;"Na jarenlang ervaring te hebben opgedaan bij toonaangevende webburo's in Friesland en in de randstad was het voor Tarek tijd een volgende stap te maken."
649	1	Ja;1/1	Ja;claimt innovatief te zijn;conceptontwikkeling: ontwikkeld onderscheidende, effectieve en aansprekende vormgeving t.b.v. verschillende doeleinde	Nee;1996 opgericht volgens LinkedIn	<20	1996	2013	2	2	0	4	

Web ID	Top25	TS-expertise	Innovatief	Startup-kenmerken	Werknemers	leeftijd o.b.v. web/LinkedIn	Wayback machine	TS-expertise	Innovatief	Startup	Validiteit-score	Type nieuwkomer
883	4	Deels;4/4;Facilerend/ Generiek, afgestemd op klanten in topsectoren	Onbevestigd	Ja;Team bestaat uit 5 medewerkers	<20	2016	2016	1	0	2	3	
596	4	Ja;4/4	Ja;claimt innovatief te zijn;"Our range of products varies from NP reducing Biopellets, Easy feed, Reef pearls, Joes Juice, Coral Protect and Ocean Prime Copepods, which we develop and manufacture in the Netherlands."	Nee;Oprichting 2001 volgens LinkedIn profiel oprichter	nvt	te oud	2004	2	2	0	4	
928	2	Ja;2/2;;+3 HTSM + AgriFood + Chemie	Ja claimt innovatief te zijn;"KULATEC is gespecialiseerd in kunststof onderdelen voor industriële installaties. Soms gebeurt dat voor een deel van het traject, bijvoorbeeld productie, soms voor turn key levering van ontwerp tot en met inbedrijfstelling en onderhoud. Voor de bouw van nieuwe instal- laties en proefinstallaties, en de revisie van bestaande installaties: KULATEC!"	Ja;2012 oprichting volgens linkedin en 2 eigenaren zichtbaar op linkedin	<20	2012	2017	2	2	2	6	Spinoff uit bedrijf;oprichter heeft volgens LinkedIn ervaring bij diverse gevestigde bedrijven en dezelfde of aanverwante sector(en)
1398	3	Deels;3/3;Facilerend/ Generiek	Onbevestigd	Ja;oprichting 2012 en 4 medewerkers volgens webpagina en zichtbaar op LinkedIn	<20	2016	2016	1	0	2	3	
885	3	Ja;3/3	"Twijfel;Onduidelijk in hoeverre bedrijf zelf ontwikkelt/innoveert;"All our programs are based on up-to-date scientific research." & "the owners and managing partners of 2M (...) had a vision of a non- standard recruitment agency where it wasn't about bulking CV's, but about providing on point quality and establishing sustainable, long term- relations. (...) They quickly built up a large network and were among the first who used social networks for their executive searches. Aiming to be pioneers with a new philosophy was one of the foundations upon which 2M International Recruitment was built." & "We'll continue to grow and develop as a company, for example by investing in new technologies such as the usage of big data and artificial networking"	Nee;"2M International Recruitment was founded in 2008 by Albert Peters & Jos Zijlstra" Oprichting 2008 en 121 medewerkers zichtbaar op LinkedIn	20+	2008	2013; op basis van gevonden website (https:// www.2m recruit- ment. com/); oude website van in- schrijving niet in gebruik volgens wayback- machine enkel in 2018 gecraw- led	2	1	0	3	
1544	3	Ja;2/3;;-1 Energie	Ja;claimt innovatief te zijn;"Deze bundeling van krachten maakt van Zorg-nv een unieke consultancy- organisatie voor de zorg en het sociale domein."	Ja;5 medewerkers volgens website;oprichting 2015 volgens LinkedIn	<20	2015	2016	2	2	2	6	Spinoff uit bedrijf; "Zorg-nv is een (openbare) maatschap van vijf maten (vennoten).Wij leveren allen een eigen, specifieke bijdrage aan onze geza- menlijke dienstverlening. Dat levert een bijzondere bundeling van krachten op; kennis, visie en ervaring die voortkomen uit de uitoefening van de volgende beroepen: huisarts, apotheker, fy- siotherapeut, consultant, projectontwikkelaar"
457	1	Ja;1/1;;+1 Creatieve Industrie	Ja;"Actief op zoek naar verbeteringen en innova- ties" advies gericht op het verbeteren en efficiënter maken van processen in de logistiek	Nee;3 medewerkers maar oprichter heeft hoofdbedrijf al in 2001 opgericht	<20	2001	2018	2	2	0	4	
900	1	Deels;0/1;bemiddeling voor o.a. ICT en Logistiek	Onbevestigd	Nee;2004 opgericht volgens LinkedIn en 3 medewerkers zichtbaar op LinkedIn	<20	2004	2005	1	0	0	1	
1230	6	Deels;6/6;Facilerend/ Generiek, afgestemd op klanten in topsectoren	Ja;claimt innovatief te zijn;"We are on a journey to Society 3.0. Change is happening faster than ever and we are keeping up. We want to give you the stories, showcases and articles to inspire you about this Society 3.0 given by people from all over the world within our network. We believe that access is more important than ownership, we don't want you just to see the story, we want you live it!"	Nee;in 32 landen 196 loca- ties met medewerkers...	20+	te oud	2005	1	2	0	3	

Web ID	Top25	TS-expertise	Innovatief	Startup-kenmerken	Werknemers	leeftijd o.b.v. web/LinkedIn	Wayback machine	TS-expertise	Innovatief	Startup	Validiteit-score	Type nieuwkomer
1342	2	Ja;2/2	Ja;claimt innovatief te zijn;"Ingenuity and flexibility are essential to us at THORCO PROJECTS and to the way that we conduct business. When meeting the individual requests of our clients, responding to market trends and when structuring our fleet, offices and organization, we continuously strive to maintain our dynamic approach of conducting business."	Nee;"Since our founding in 2003, we have rapidly established ourselves as an industry leader in project shipping."	nvt	2003	2017	2	2	0	4	
496	4	Ja;4/4	Ja;claimt innovatief te zijn;"COSMONiO's office is on the internet, since we always aim to work with the brightest minds in the world regardless of their location. Still, we need labs for our research and development activities, so we founded CORTEXLAB, HEALTHLAB and LIFELAB. Both of these labs are positioned in the heart of highly innovative campuses, where science, technology and creativity blend."	Ja;Team bestaat uit 11 medewerkers	<20	2013	2013; 2010 crawled, website had toen ander doeleinde	2	2	2	6	"Incumbent-backed entrant; ""COSMONiO is part of NVIDIA INCEPTION PROGRAM. Inception nurtures dedicated and exceptional startups who are revolutionizing industries with advances in AI and data science.""
1047	2	Ja;2/2;;+1 Logistiek	Ja;claimt inovatief te zijn;"Onze Business Unit Concepts houdt zich bezig met nieuwe technieken en 'new business' en/of startups. Hierbij gaat het vooral om technieken en/of ideeën die aan het begin staan van brede acceptatie. Om deze technieken en ideeën een kans te geven is een unieke samenwerking gestart tussen: Bedrijfsleven – Opleiding – IT bedrijven – Overheid" & "Naast al onze IT producten en diensten ondersteunen wij bedrijven voor het opzetten van nieuwe businessideeën o.b.v. door ons ontwikkelde diensten en / of producten" & "Ontwikkeling van software- en marktconcepten"	Nee;2008 opgericht n 14 medewerkers zichtbaar op linkedin	<20	2008	2010	2	2	0	4	
1172	3	Ja;3/3	Ja;"Het ontwikkelen van inventieve software-producten houdt ons in ontwikkeling, waardoor we ons blijven verbeteren..."	Nee;"AL 12 JAAR BETROKEN BIJ DE ICT-AMBITIE VAN TRANSPORTEURS EN TECHNEUTEN" LinkedIn info oprichting 2007 en 2 medewerkers zichtbaar op LinkedIn	<20	2007	2012	2	2	0	4	Spinoff uit bedrijf;"Reinder.NET. Maintenance 2012, de opvolger van de ServiceBase, is in Beta versie beschikbaar"
1434	2	Ja;2/2;;+1 Logistiek	Twijfel;onduidelijk in hoeverre bedrijf zelf ontwikkelt/innoveert, lijkt vooral bestaande diensten en producten te vermarkten	Nee;2009 opgericht en 53 medewerkers zichtbaar op LinkedIn	20+	2009	2011	2	1	0	3	
72	3	Ja;3/3	Ja;claimt innovatief te zijn;"Elaborate process know-how and extensive technological experience provide us with a leading role advantage in the precise electrochemical machining technology's future development."& "Those process promotion activities, together with our extensive R&D knowledge capabilities have placed us at a leading thought leadership position within the pECM market. We recognize the responsibility of maintaining our knowledge authority role, and thus strive to keep on offering superior products and services to our current and potential clients."	Nee;2003 opgericht	nvt	te oud	2003	2	2	0	4	
307	9	Deels;9/9;Faciliterend/Generiek, afgestemd op klanten in topsectoren	Onbevestigd	Onbevestigd	<20	2013	2013	1	0	2	3	

Web ID	Top25	TS-expertise	Innovatief	Startup-kenmerken	Werknemers	leeftijd o.b.v. web/LinkedIn	Wayback machine	TS-expertise	Innovatief Startup	Validiteit-score	Type nieuwkomer	
496	4	Ja;4/4	Ja;claimt innovatief te zijn;"COSMONIO's office is on the internet, since we always aim to work with the brightest minds in the world regardless of their location. Still, we need labs for our research and development activities, so we founded CORTEXLAB , HEALTHLAB and LIFELAB. Both of these labs are positioned in the heart of highly innovative campuses, where science, technology and creativity blend."	Ja;Team bestaat uit 11 medewerkers	<20	2013	2013; 2010 crawled, website had toen ander doeleinde	2	2	2	6	"Incumbent-backed entrant; ""COSMONIO is part of NVIDIA INCEPTION PROGRAM. Inception nurtures dedicated and exceptional startups who are revolutionizing industries with advances in AI and data science."" "
846	3	Ja;3/3	Ja;claimt innovatief te zijn;"Innovative Technical solutions" & "Vanuit Drachten ontwikkelt ITS innovatieve en technologisch geavanceerde oplossingen voor bedrijven die zelf onvoldoende ontwikkelcapaciteit of kennis in huis hebben."	Ja;LinkedIn:" Year founded 2012 (...) Company size 2-10 employees"	<20	2012	2012	2	2	2	6	Spinoff uit bedrijf;"With 22 years experience in defense (marines and airforce) I am an initiative and assertive person, a "quick thinker", able to analyse company operations, tactics and strategies because of my broad experience at all company levels. "
110	5	Nee;0/6;Groothandel	Onbevestigd	Ja	<20	2015	2015	0	0	2	2	
1095	5	Ja;4/5;;;-1 AgriFood	Ja;claimt innovatief te zijn;"JONG, INNOVATIEF EN BETROUWBAAR"	Ja;2015 opgericht en 2 medewerkers zichtbaar op LinkedIn	<20	2017	2017	2	2	2	6	Diversifying entrant; LinkedInprofielen van de oprichters met werkervaringen in enkele marketing en ICT.
1166	1	Nee;0/1;Groothandel	Onbevestigd	Ja;2014 opgericht volgens LinkedIn	<20	2014	2015	0	0	2	2	
1569	3	Deels;3/3;vooral gericht op uitvoering en informeren zonder zelf te innoveren	Onbevestigd	Onbevestigd	<20	2014; 2013 crawled maar geen info/foutmelding	2014; 2013 crawled maar geen info/foutmelding	1	0	2	3	
390	2	Ja;2/2	"Ja;claimt innovatief te zijn;"BoekTweePuntNul is geen traditionele uitgeverij maar een auteurscollectief waar inmiddels meer dan 200 co-auteurs aan verbonden zijn."" Brengt boeken uit waarin experts de mogelijkheden van sociale media, apps en andere digitale toepassingen binnen het onderwijs uiteenzetten. "	Ja;2011 oprichting en 7 medewerkers zichtbaar op LinkedIn	<20	2011	2011	2	2	2	6	Spinoff uit bedrijf;cocreatie van initiatiefnemers met werkervaring in de ICT, onderwijs en schrijven/ auteurschap
496	4	Ja;4/4	Ja;claimt innovatief te zijn;"COSMONIO's office is on the internet, since we always aim to work with the brightest minds in the world regardless of their location. Still, we need labs for our research and development activities, so we founded CORTEXLAB , HEALTHLAB and LIFELAB. Both of these labs are positioned in the heart of highly innovative campuses, where science, technology and creativity blend."	Ja;Team bestaat uit 11 medewerkers	<20	2013	2013; 2010 crawled, website had toen ander doeleinde	2	2	2	6	"Incumbent-backed entrant; ""COSMONIO is part of NVIDIA INCEPTION PROGRAM. Inception nurtures dedicated and exceptional startups who are revolutionizing industries with advances in AI and data science."" "
714	4	Ja;4/4	Ja;claimt innovatief te zijn;Innovatieaward winner	Nee;2003 ingeschreven	<20	te oud	2003	2	2	0	4	"En zo ontstond in 1999 de naam Grendel Games. In 2002 werd Grendel Games officieel ingeschreven bij de Kamer van Koophandel."
1230	6	Deels;6/6;Faciliterend/Generiek, afgestemd op klanten in topsectoren	Ja;claimt innovatief te zijn;"We are on a journey to Society 3.0. Change is happening faster than ever and we are keeping up. We want to give you the stories, showcases and articles to inspire you about this Society 3.0 given by people from all over the world within our network. We believe that access is more important than ownership, we don't want you just to see the story, we want you live it!"	Nee;in 32 landen 196 locaties met medewerkers...	20+	te oud	2005	1	2	0	3	

Web ID	Top25	TS-expertise	Innovatief	Startup-kenmerken	Werknemers	leeftijd o.b.v. web/LinkedIn	Wayback machine	TS-expertise	Innovatief	Startup	Validiteit-score	Type nieuwkomer
1290	2	Ja;2/2	Ja;claimt innovatief te zijn;"Als leider in de MODX-community is innoveren een levensbehoefte." & "Al meer dan 13 jaar zijn we MODX-partner. We ontwikkelen met en aan deze open source techniek."	Nee;2004 opgericht volgens linkedin en 25 medewerkers zichtbaar op LinkedIn e 24 op de website	20-25	2004	2011; 2004 crawled maar met foutmelding	2	2	0	4	
1551	3	Ja;3/3	"Ja;claimt innovatief te zijn;""We always keep in mind that our innovative solutions must be robust, effective, user friendly, scalable and must have a very limited integration time in daily work flows. This is key for us! Together with our inhouse development team and trusted partners we are busy with the development of new solutions and to improve current applications. Besides this we take all feedback and input given by our clients very seriously and we add new functionality to our current solutions on a regular base. Our VR solutions are being used globally. Therefore, we have assigned a network of sales consultants & highly skilled trainers to stay in close contact and being your nearby partner."" & ""In close collaboration with our professional partners we developed 2 solutions: VRiend and VR-Macy""	Ja;oprichting 2017 volgens linkedin en minimaal 2 medewerkers, althans website redirect naar nieuw bedrijf Horus VR (https://vr-iend.com/), oorspronkelijke bedrijf en website	<20	2017	2017; oorspronkelijke website sinds 2009 crawled	2	2	2	6	Spinoff uit bedrijf; oprichters met sectorgerelateerd werkervaring bij inmiddels gevestigde bedrijven
1572	3	Ja;3/3	Ja;claimt innovatief te zijn;"wij zijn er om het online landschap elke dag een stukje mooier en slimmer te maken. Wij bedenken en ontwikkelen strategische design producten die zich onderscheiden van de massa."	Nee;opgericht in 2004 en 23 medewerkers	20+	2004	2004	2	2	0	4	
1364	6	Deels;6/6;Faciliterend/Generiek	Ja;claimt innovatief te zijn;"We continue our quest for niche markets"	Ja;oprichting 2012 en 6 medewerkers zichtbaar op LinkedIn. Grotere partners wel al veel langer actief (spinoff)	<20	2012	2013; 2012 informeel	1	2	2	5	Incumbent-backed entrant; Torrgas is a Dutch, privately funded, clean technology company which has developed a state of the art process for the production of Syngas from torrefied biomass and subsequent Syngas upgrading.
307	9	Deels;9/9;Faciliterend/Generiek, afgestemd op klanten in topsectoren	Onbevestigd	Onbevestigd	<20	2013	2013	1	0	2	3	
1049	3	Ja;3/3	Twijfel;Produceert (ver) nieuwe(nde) producten	Nee;"In business since 1907, the Nielsen-Massey brand has built an enviable reputation for the quality and consistency of its pure vanillas."	nvt	te oud	2001	2	1	0	3	
1398	3	Deels;3/3;Faciliterend/Generiek	Onbevestigd	Ja;oprichting 2012 en 4 medewerkers volgens webpagina en zichtbaar op LinkedIn	<20	2016	2016	1	0	2	3	
1433	1	Ja;1/1;+1 Creatieve Industrie	Ja;claimt innovatief te zijn;"Wat begon met expertise op het gebied van veevoeding is uitgegroeid tot een adviesbureau die het vernieuwende adviesconcept IndeXboeren heeft ontwikkeld. Met dit adviesconcept, waarin de bodem, voer en koe centraal staan, ondersteunen wij de agrarisch ondernemer."	Nee;"In 1987 is het VIB (Vee- en Tuinbouwkundig Ingenieursbureau) begonnen met haar werkzaamheden als onafhankelijk adviesbureau."14 medewerkers zichtbaar op LinkedIn en 13 medewerkers volgens website	<20	1987	2009	2	2	0	4	

Web ID	Top25	TS-expertise	Innovatief	Startup-kenmerken	Werknemers	leeftijd, o.b.v. web/LinkedIn	Wayback machine	TS-expertise	Innovatief Startup	Validiteit-score	Type nieuwkomer	
601	2	Ja;2/2;;+2 Energie + Chemie	Ja;claimt innovatief te zijn;"Econvert Water & Energy B.V. designs and implements technology for the treatment of industrial wastewater & biogas. Because of the unique settling systems our reactors can handle a wide variety and different concentrations of waste water. Through of our lean and mean designs, anaerobic technology is now available for a large number of industries." & "Econvert is very distinctive thanks to our thorough process knowledge and we the possibility to offer tailor-made solutions such as rental installations."	Ja;2012 opgericht n 24 medewerkers zichtbaar op LinkedIn	20-25	2012	2011	2	2	2	6	Spinoff uit bedrijf; met behulp van wayback machine en linkedin blijkt dat bedrijf is ontstaan uit eerdere oprichting van Econvert VOF in 2005, Econvert Climate & Energy and Econvert Water & Energy zijn hier later uit voortgekomen.
840	4	Ja;3/4;;;-1 Water	Ja;claimt innovatief te zijn;"In 2012 we decided to focus our Research & Development and portfolio extension on Women's Health. This resulted in several patents and patent applications in the field of prognostic assays to determine the risk and pregnancy related disorders."&"Unique and ready-to-use antibodies and complete innovative diagnostic kits developed for all currently available flow cytometers."	Nee;opgericht in 1987	nvt	te oud	1999	2	2	0	4	;"It was founded in cooperation with the leading University of Groningen, in the north of The Netherlands in 1987."
304	5	Ja;5/5	Ja;claimt innovatief te zijn;"Arox staat voor innovatie. Met onze nieuwste innovaties zorgen we dat bedrijven klaar voor de toekomst zijn. De techniek van morgen, vindt u vandaag al bij ons. Neem een kijkje op de pagina van Arox Smart Technologies en leer onze autonome 'tel-drone' en de smartglasses kennen."	Nee;op basis van website al sinds 2008 actief	<20	te oud	2008; 2005 crawled maar geen info/foutmelding	2	2	0	4	
496	4	Ja;4/4	Ja;claimt innovatief te zijn;"COSMONIO's office is on the internet, since we always aim to work with the brightest minds in the world regardless of their location. Still, we need labs for our research and development activities, so we founded CORTEXLAB , HEALTHLAB and LIFELAB. Both of these labs are positioned in the heart of highly innovative campuses, where science, technology and creativity blend."	Ja;Team bestaat uit 11 medewerkers	<20	2013	2013; 2010 crawled, website had toen ander doeleinde	2	2	2	6	"Incumbent-backed entrant; ""COSMONIO is part of NVIDIA INCEPTION PROGRAM. Inception nurtures dedicated and exceptional startups who are revolutionizing industries with advances in AI and data science.""
1230	6	Deels;6/6;Faciliterend/Generiek, afgestemd op klanten in topsectoren	Ja;claimt innovatief te zijn;"We are on a journey to Society 3.0. Change is happening faster than ever and we are keeping up. We want to give you the stories, showcases and articles to inspire you about this Society 3.0 given by people from all over the world within our network. We believe that access is more important than ownership, we don't want you just to see the story, we want you live it!"	Nee;in 32 landen 196 locaties met medewerkers...	20+	te oud	2005	1	2	0	3	
786	2	Ja;2/2	Ja;claimt innovatief te zijn;Benoemt niet heel expliciet de innovatie maar wel de nieuwe producten en de wetenschappelijke expertise die ingezet wordt voor de ontwikkeling en productie: "New! 2d trnsfer system. Designed and assembled at HQ Graphene. (...) This manual transfer system is designed to accurately place a stamp on a substrate, which can be moved relative to each other not only in the x-, y- and z-directions, but also by tilting and rotating. This allows the highest degree of freedom in the alignment of crystal flakes when fabricating heterostructures. The substrate is fixated with vacuum and can be heated during the process. The compact design allows for easy placement inside a glove box, providing the option to work with air sensitive materials." & "Our scientific staff has the expertise in the production and characterization of large size and high purity single crystals. The crystals are characterized with X-ray diffraction (single crystal XRD and powder XRD), Energy-dispersive X-ray spectroscopy (EDX), Raman and additionally with Hall measurements. We can grow all kinds of two dimensional crystals:"	Ja;oprichting 2012	<20	2012	2013	2	2	2	6	Spinoff uit onderzoeksinstelling; Volgens LinkedIn: PhD in Applied Physics: "As a PhD student I performed spin transport measurements in graphene and carbon nanotubes. & R&D Experimental Physicist bij MAPPER lithography (gevestigd bedrijf 200+FTE)

Web ID	Top25	TS-expertise	Innovatief	Startup-kenmerken	Werknemers	leeftijd o.b.v. web/LinkedIn	Wayback machine	TS-expertise	Innovatief	Startup	Validiteit-score	Type nieuwkomer
76	2	Nee;0/2;Groothandel	Onbevestigd	Nee;bedrijf niet vermoedelijk niet langer levensvatbaar, website blijkt niet up to date en op linkedin alleen maar voormalige werknemers te vinden waaronder de oprichter. 2009 opgericht en 2011 niet langer actief volgens linkedin informatie.	<20	2009	2010	0	0	0	0	
628	2	Nee;0/2;Groothandel	Onbevestigd	Nee;bedrijf niet vermoedelijk niet langer levensvatbaar, website blijkt niet up to date en op linkedin alleen maar voormalige werknemers te vinden waaronder de oprichter. 2009 opgericht en 2011 niet langer actief volgens linkedin informatie.	<20	2009	2010	0	0	0	0	
120	2	Deels;2/2; Faciliterend/Generiek	Twijfel;Platform, Incubator, Uitgever	Ja;2017 opgericht volgens linkedin, 4 medewerkers volgens website en al sinds crawl in 2013 operationeel volgens waybackmachine	<20	2017	2014	1	1	2	4	
246	3	Ja;3/3	Ja;claimt innovatief te zijn;"Bij 8D Games wordt veel gebruik van brainstormsessies om ideeën te bedenken (...) om creatieve oplossingen te verzinnen."	Ja;oprichting 2012 en 15 medewerkers zichtbaar op LinkedIn	<20	2012	2013	2	2	2	6	Diversifying entrant; Oprichter volgens linkedin voorheen met werkervaring in andere sector(en)
416	1	Ja;1/1	Onbevestigd	Nee;2009 opgericht 1 medewerker zichtbaar op linkedin (bedrijfsinformatie geeft 2-10 medewerkers aan)	<20	2009	2010	2	0	0	2	
1300	3	Ja;3/3	Ja;Ja;claimt innovatief te zijn;Conceptontwikkeling "Al onze producten hebben gemeen dat ze unieke visuals"	Ja;Oprichting 2013 en team bestaat uit 2 medewerkers	<20	2013	2015; 2009 crawled maar geen info/foutmelding	2	2	2	6	Spinoff uit bedrijf; oprichters met sectorgerelateerd werkervaring bij gevestigde bedrijven
1398	3	Deels;3/3;Faciliterend/Generiek	Onbevestigd	Ja;oprichting 2012 en 4 medewerkers volgens webpagina en zichtbaar op LinkedIn	<20	2016	2016	1	0	2	3	

B2 Opschonen Externe Data

In deze bijlage is beknopt beschreven hoe er is omgegaan met herindelingen van buurt- en gemeentegrenzen en daarbinnen de ontbrekende waarden. Alle externe data waren afkomstig van de StatLinepublicaties Kerncijfers wijken en buurten van het CBS: <https://www.cbs.nl/nl-nl/dossier/nederland-regionaal/wijk-en-buurtstatistieken/kerncijfers-wijken-en-buurten-2004-2018>

Uit tabel B2.1 is af te lezen welke variabelen op welk schaalniveau beschikbaar waren en dus gebruikt konden worden. Voor alle variabele zijn de meest recente gegevens gebruikt, de jaartallen zijn af te lezen in de tabel. Variabelen die alleen voor de jaren 2016 of 2017 beschikbaar waren moesten worden omgezet naar de buurt- en gemeente indeling van 2018 omdat in Noord-Nederland verschillende gemeentelijke herindelingen sinds 1 januari 2018 van kracht zijn gegaan:

- De gemeenten Franekeradeel, het Bildt, Menameradiel en Littenseradiel zijn opgeheven en de nieuwe gemeente Waadhoeke is ingesteld, bestaande uit het grondgebied van Franekeradeel, het Bildt, Menameradiel en een deel van Littenseradiel. De andere delen van Littenseradiel zijn heringedeeld met Leeuwarden en Súdwest-Fryslân.
- De gemeenten Leeuwarderadeel is opgeheven

en toegevoegd aan de gemeente Leeuwarden.

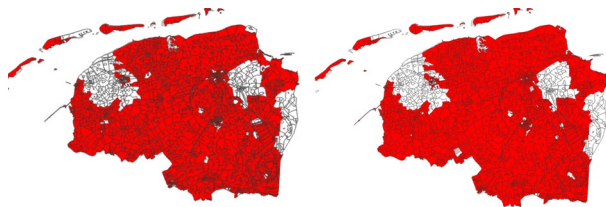
- De gemeente Vlagtwedde fuseerde met de gemeente Bellingwedde tot de nieuwe gemeente Westerwolde.
- De gemeente Hoogezand-Sappemeer fuseerde met de gemeenten Slochteren en Menterwolde tot de nieuwe gemeente Midden-Groningen

Data m.b.t. hoogopgeleiden uit 2017 kon op gemeenteniveau worden geconverteerd. Bij samenvoegingen van gemeenten konden absolute aantallen worden opgeteld terwijl relatieve aantallen en opsplitsingen van gebieden opnieuw moesten worden berekend.

Bij sommige gegevens op buurtniveau afkomstig van het CBS, zoals het aantal bedrijfsregistraties en de nabijheidsvariabelen per buurt, waren de meest recente gegevens alleen beschikbaar voor het jaar 2016. In figuur B2.1 is te zien bij welke buurten er conversieproblemen optraden. Er was

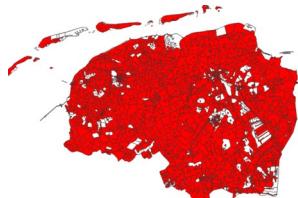
Tabel B2.1 Beschikbaarheid Data.

Data	kleinst beschikbare schaalniveau	meetmoment
Query resultaten	Coördinaten	2018
Totaal aantal inschrijvingen bedrijven	Gemeente incl. totaal NL	2018 (zelfde als queryresultaten)
Onderwijsinstelling variabele (geïnterviewde hogescholen en universiteiten)	Coördinaten	2012-2018
CBS-gegevens:		
Bevolking etniciteit, leeftijd, omgevingsadressendichtheid	Buurt	2018
Gemiddelde WOZ waarde	Buurt	2017
Hoogopgeleiden	Gemeente	2017
Nabijheid voorzieningen en de dichtheid van voorzieningen (inclusief infrastructuur)	Buurt	2016
Aantal bedrijven	Buurt	2016



Figuur B2.1 Screenshots van ontbrekende waarde nabijheidsvariabelen uit 2016 voor de buurtindeling van 2018. De dekking van de variabelen Afstand oprit hoofdverkeersweg per buurt (linksboven) en Bedrijfsvestigingen totaal per buurt (rechtsboven) is rood gekleurd. De witte gebieden ontbreken. Voor de variabelen Afstand belangrijk overstapstation en aantal cafés binnen 1 km ontbreken dezelfde gebieden.

echter geen goed werkende koppeltabel beschikbaar om dit probleem op te lossen. Als alternatieve oplossing zijn de relevante statistieken verbonden aan verouderde buurtcentroids (geometrische zwaartepunten) gekoppeld aan de dichtstbijzijnde buurtcentroids van de nieuwe indeling.

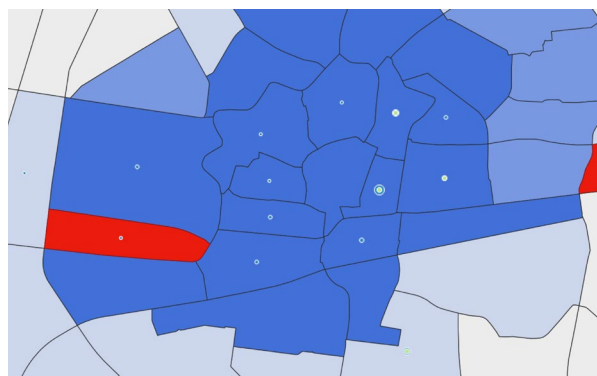


Figuur B2.2 Screenshots van ontbrekende waarde voor westerse allochtonen per buurt. De dekking van de variabelen is rood gekleurd. De witte gebieden ontbreken.

Nadat het probleem m.b.t. herindelingen was opgelost, moest worden gekeken naar ontbrekende waarden in de data. Hier is per variabele anders mee omgegaan. Bij sommige gegevens, zoals de gemiddelde WOZ-waarde in een buurt, waren dusdanig veel ontbrekende waarde dat ze niet konden worden betrokken in de analyse. Andere gegevens, zoals bijvoorbeeld het aantal westerse allochtonen per buurt (figuur B.2.2) waren minder problematisch. In het laatste geval is nader bekeken waar en waarom er waarden missen. Bij westerse allochtonen had dit vaak te maken met de kleine aantallen

inwoners in de buurt. In deze gevallen zijn ontbrekende waarde omgezet naar 0.

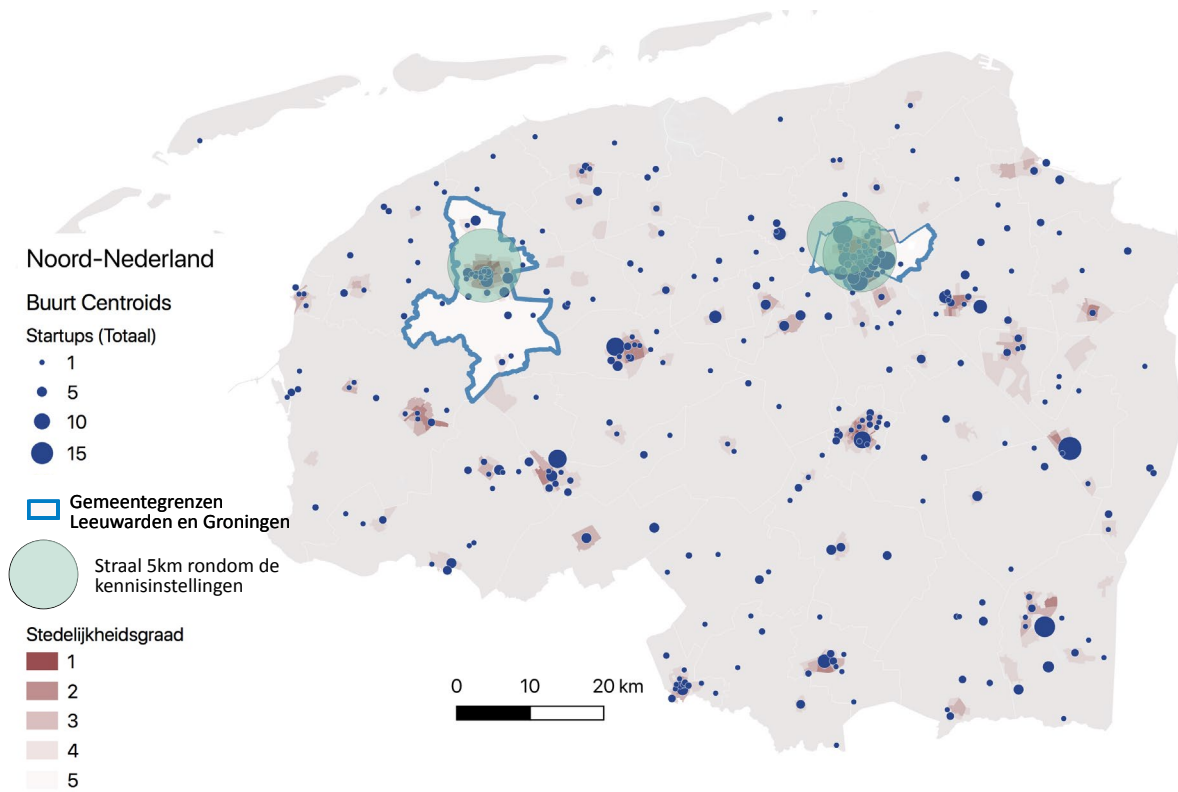
Bij ontbrekende waarde m.b.t. aantallen cafés binnen een straal van 1km is voor enkele buurten handmatig bijgeschat. In figuur B2.3 is een voorbeeld te zien van een situatie in de stad Leeuwarden waar is bijgeschat in buurten waar hoogscorend startupbedrijven voorkomen. In het rode vlak links in de figuur komt een hoogscorend startupbedrijf voor. Aangezien in die buurt het aantal cafés ontbreekt is gekeken naar de waarden van de omliggende buurten. Bij de omliggende buurten waren aantallen cafés in een straal van 1, 3 en 5km bekend, daarnaast is op google maps gekeken naar bijzonderheden in de buurt. Het was echter niet duidelijk waarom voor sommige buurten waarden ontbreken. Op basis van Google Maps en de omliggende buurten is om redenen van efficiëntie alleen het aantal cafés bijgeschat voor buurten met hoogscorende startups. rige ontbrekende variabelen zoals het aantal bedrijven in 2016 was geen simpele oplossing voor handen. Toch is met behulp van de besproken maatregelen het aantal cases (buurten) met ontbrekende waarde beperkt tot een minimum.



Figuur B2.3 Screenshots van aantallen cafés per buurt met buurtcentroids van hoogscorende startup bedrijven in een stedelijk deel van het onderzoeksgebied. Donkerblauwe vlakken corresponderen met een hoog aantal cafés terwijl de rode vlakken ontbrekende waarden betreffen. De stippen zijn bedrijven met een hoge relevantiescore (blauw = top25% score en groen = top25% score en ook een hoge valditeitscore van vijf of zes bij de steekproef).

B3 Overige Figuren en Tabellen

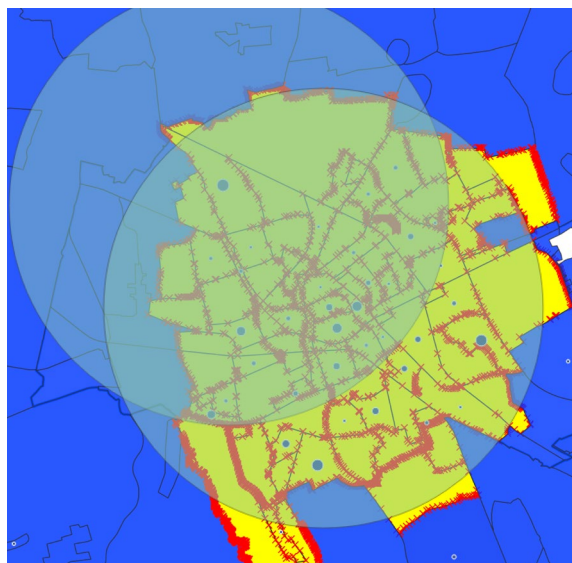
In deze bijlage zijn overige figuren en tabellen opgenomen en in sommige gevallen nader toegelicht.



Figuur B3.1 Buurtcentroids met aantallen hoogscorende Startups ('Top25') in relatie tot de kennisinstellingen van Leeuwarden en Groningen. Bronnen: CBS (2018) en het Kadaster (2018).

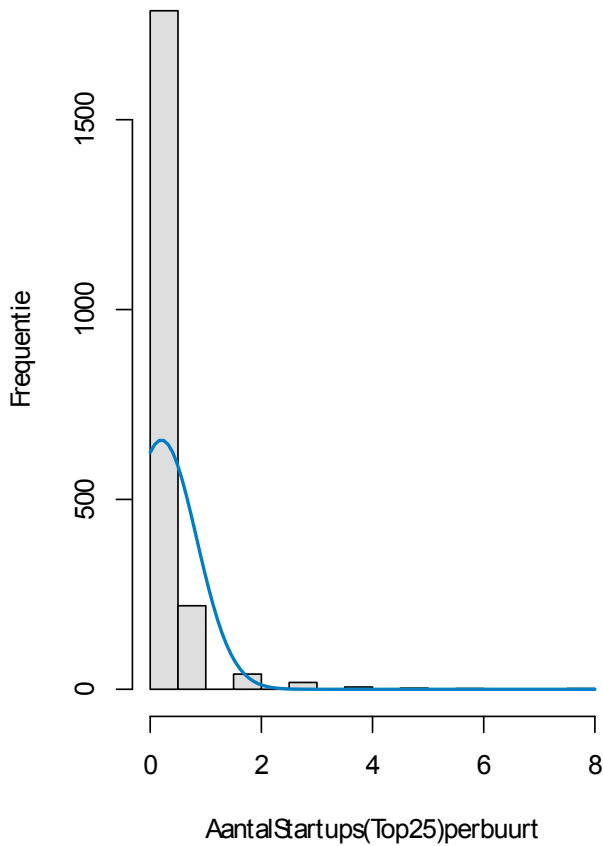
B3.1 Variabele Kennisinstelling

De groene cirkels in figuur B.3.1 zijn in een straal van vijf kilometer rondom de adressen van de HBO en WO kennisinstellingen (Dienst Uitvoering Onderwijs [DUO], 2019) getrokken. Het gaat om een hogeschool in Leeuwarden (de meest linkse groene cirkel) en een hogeschool en universiteit in Groningen (de meest rechtse groene cirkels). De straal van vijf kilometer doet er niet veel toe, aangezien binnen deze cirkels alleen buurten geselecteerd zijn die binnen de bebouwde kom (gebieden een verhoogde stedelijkheidsgraad van een tot vier) vallen. In het geval van de kennisinstellingen van Groningen is bijvoorbeeld in figuur B.3.2 te zien dat de noordelijke kennisinstelling (Hanze Hogeschool) nabij de stadsrand ligt en dus weinig invloed heeft bovenop de geselecteerde buurten binnen de bebouwde kom rondom de Universiteit Groningen.

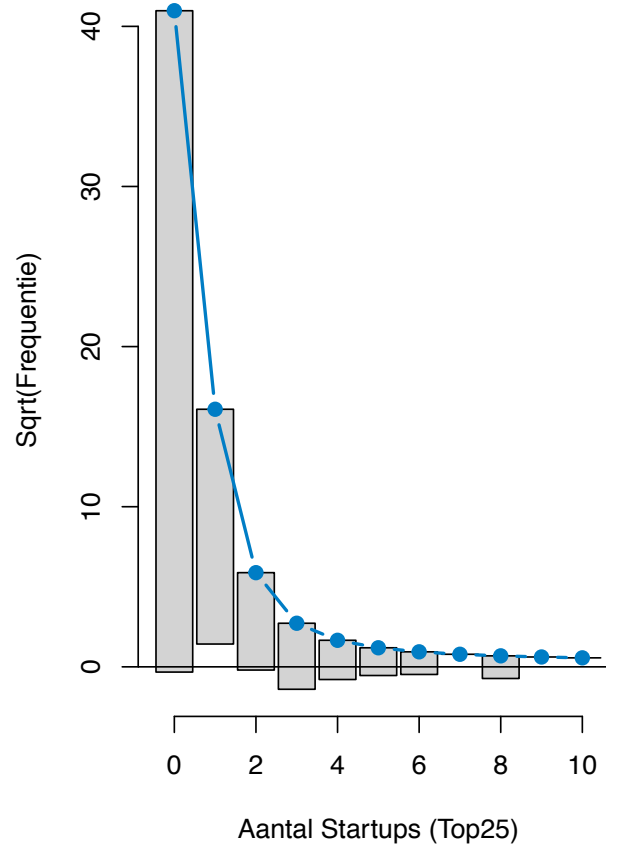


Figuur B3.2 Screenshots van de buurtselectie rondom de kennisinstellingen in Groningen.

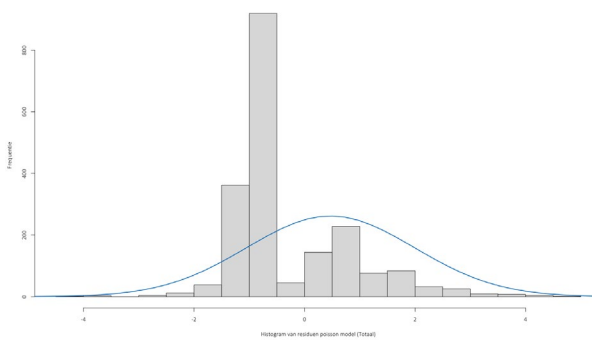
B3.2 Figuren t.b.v. de poissonverdeling



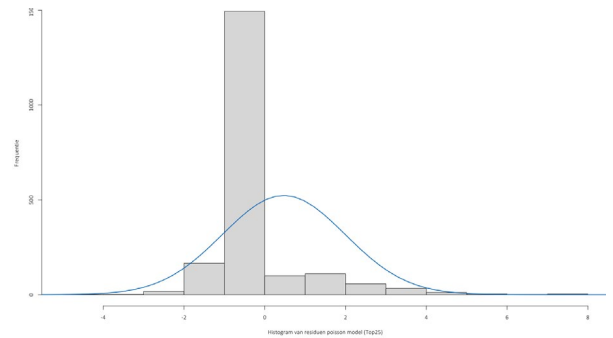
Figuur B.3.3 Histogram met normaal curve van het aantal startups per buurt (Top25) per buurt.



Figuur B.3.4 Rootogram (hangend) aanpassing van het poisson-model aan de simulatie. De blauwe lijn is de voorspelling van het poisson model, de hangende grijze balken zijn de daadwerkelijke observaties. Als de grijze balken boven de nul-lijn zweeft is er sprake van overdispersie en als deze er onder de nul-lijn door gaat is er sprake van onderdispersie.



Figuur 3.5 Histogram met normale verdeling van het aantal startups (Totaal) per buurt.



Figuur 3.6 Histogram met normale verdeling van het aantal startups (Top25) per buurt.

B3.3 Thematische kaarten

De volgende figuren tonen alle Top25-Startups naar thema, geplot met de stedelijkheidsgraad op de achtergrond.

Startups (Top25)

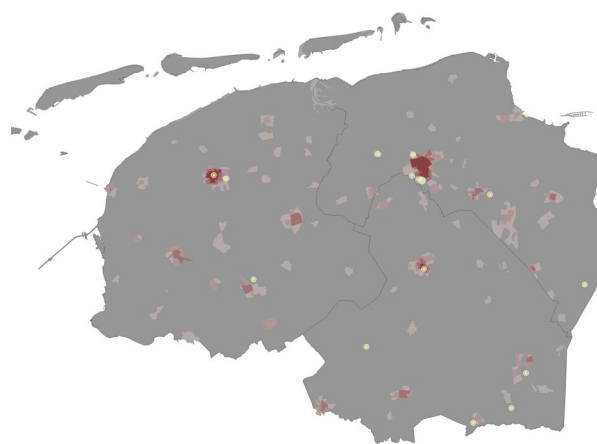
- 1
- 2
- 3
- 4
- 5

Mate van stedelijkheid

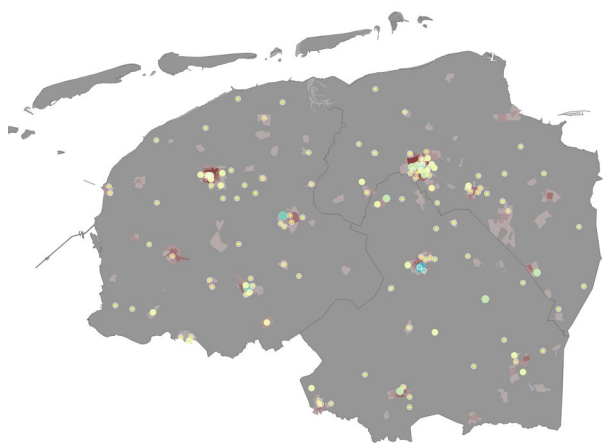
- 1
- 2
- 3
- 4
- 5



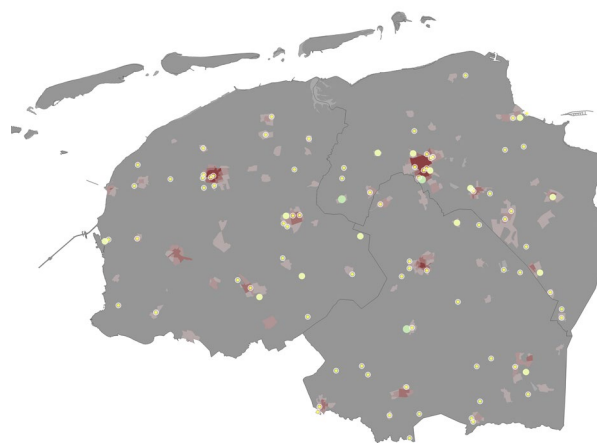
Figuur B.3.8 *Agri & Food*



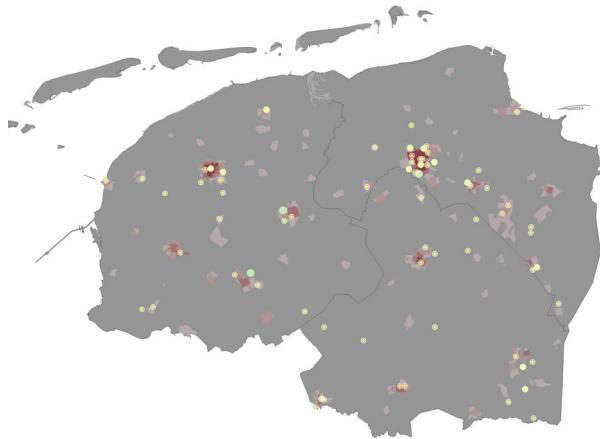
Figuur B.3.9 *Chemie*



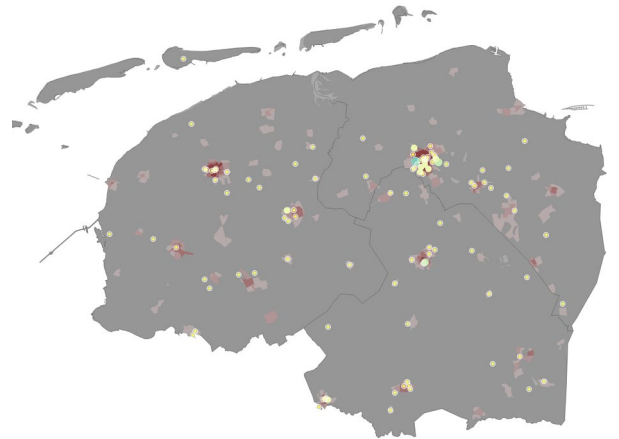
Figuur B.3.10 *Creatieve Industrie*



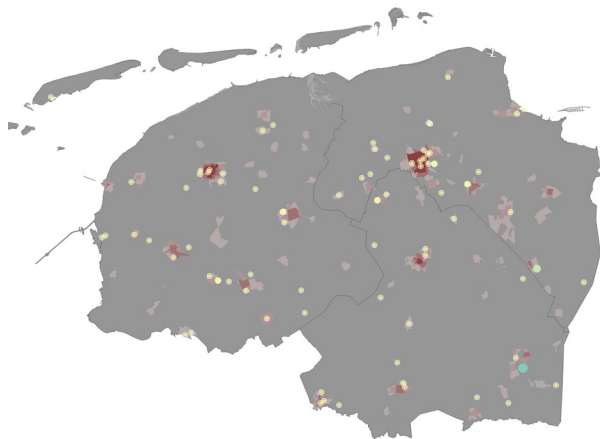
Figuur B.3.11 *Energie*



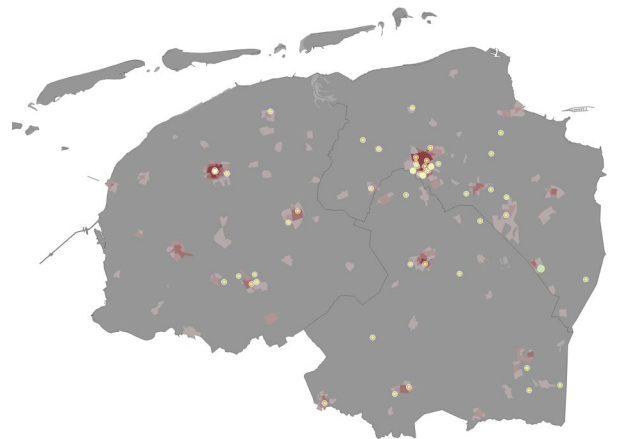
Figuur B.3.12 *HTSM*



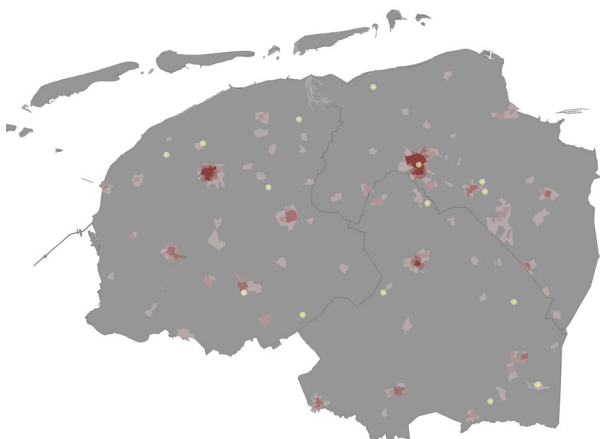
Figuur B.3.13 *ICT*



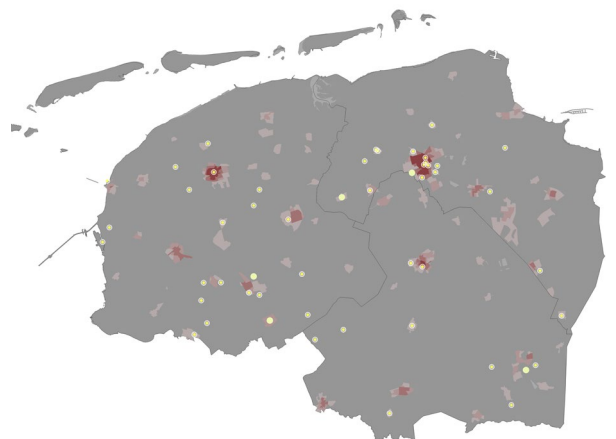
Figuur B.3.14 *LS&H*



Figuur B.3.15 *Logistiek*



Figuur B.3.16 *Tuinbouw*



Figuur B.3.17 *Water*

B4 Ruimtelijke Statistiek

In deze bijlage worden verschillende toelichtingen op methoden voor verschillende facetten van de ruimtelijke statistische analyse beschreven. De berekeningen zijn gemaakt met het opensource software programma 'R'.

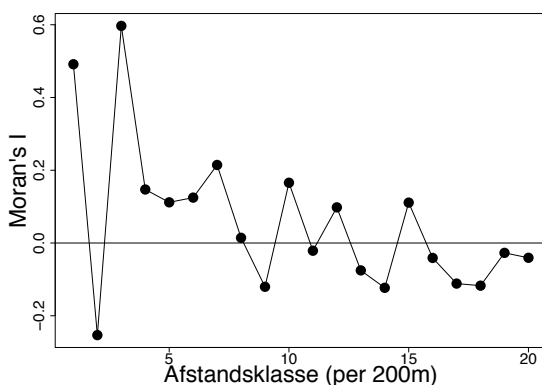
B4.1 Ruimtelijke Autocorrelatie

Ruimtelijke autocorrelatie meet in dit onderzoek de correlatie van de buurtvariabelen in Noord-Nederland met het ruimtelijke gewogen gemiddelde van nabij gelegen buurtvariabelen. Ruimtelijke autocorrelatie kan positief of negatief zijn. Positieve ruimtelijke autocorrelatie treedt op wanneer soortgelijke waarden dicht bij elkaar voorkomen. Negatieve ruimtelijke autocorrelatie treedt op wanneer ongelijksoortige waarden in de buurt van elkaar voorkomen.

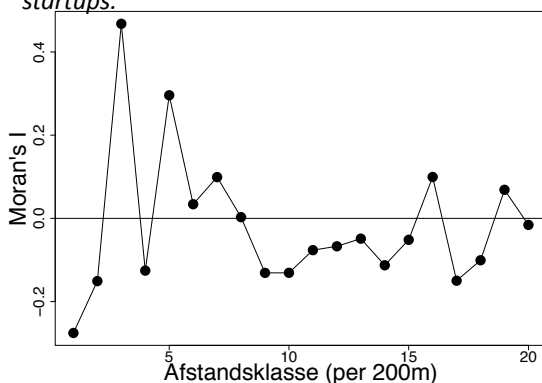
De gehanteerde maat voor ruimtelijke autocorrelatie is tussen de zwaartepunten (*centroids* met aantallen startups) van de buurten die vlak bij elkaar liggen en dat deze correlatie stabiliseert rondom nul bij een afstand van ongeveer 3 tot 4km.

i. De nulhypothese van de Global Moran's *I* stelt dat het aantal startups willekeurig verdeeld is over het onderzoeksgebied.

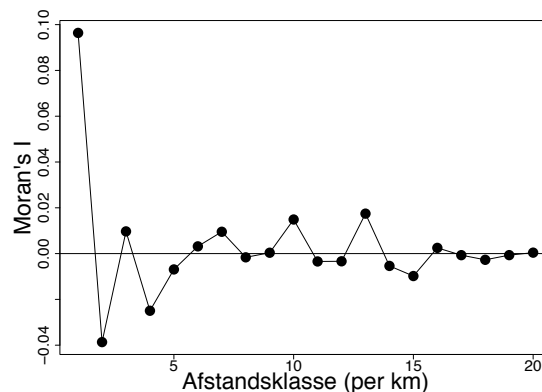
Het Moran correlogram (figuur. B.4.1 tot en met B.4.4) berekent de *I*-waarde met verschillende afstandsintervallen (Bjørnstad, Ims, & Lambin, 1999; Bjørnstad & Falck, 2001; Epperson, 1993). De plots geven aan hoe geconcentreerd of gedistribueerd de ruimtelijke autocorrelatie is. De Moran's *I* in de plots laten zien dat er positieve ruimtelijke correlatie is tussen de zwaartepunten (*centroids* met aantallen startups) van de buurten die vlak bij elkaar liggen en dat deze correlatie stabiliseert rondom nul bij een afstand van ongeveer 3 tot 4km.



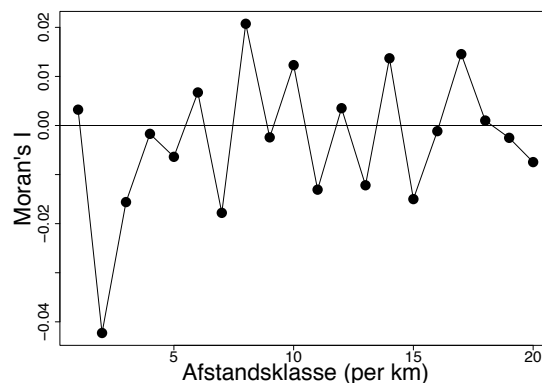
Figuur B.4.1 Moran's *I* voor het totaal aantal startups.



Figuur B.4.3 Moran's *I* voor het aantal Top25 startups.



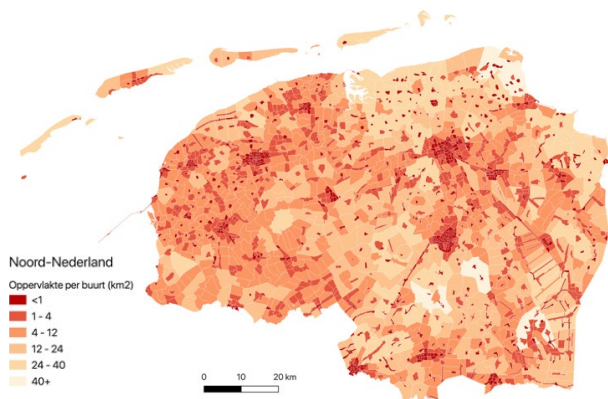
Figuur B.4.2 Zelfde als B.4.1, ander interval.



Figuur B.4.4 Zelfde als B.4.3, ander interval.

B.4.4.1 Conceptualisatie ruimtelijke relaties

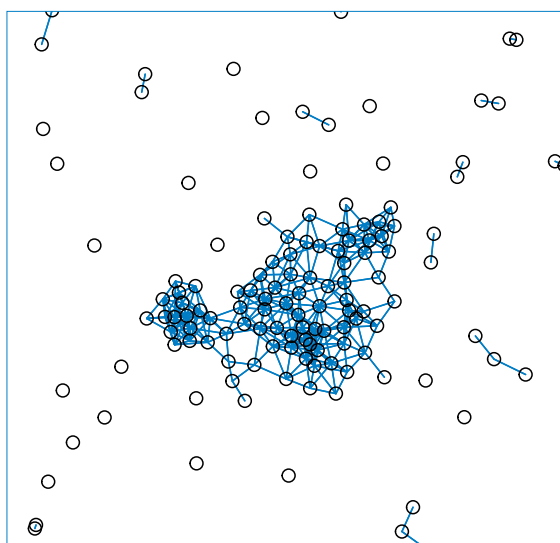
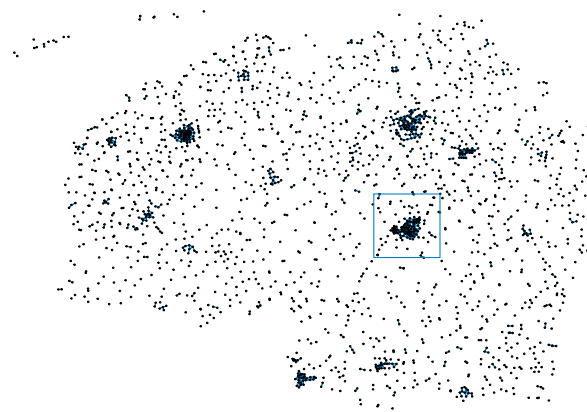
In totaal werden 1994 buurten (2078 – 86 ontbrekende waarden) met een gemiddelde oppervlakte van 4 vierkante kilometer betrokken in de analyse. Figuur B.4.5 laat zien dat de buurten verschillend van omvang zijn. Veel buurten in de stedelijke gebieden zijn kleiner dan 1 vierkante kilometer en er zitten ook een aantal extreem grote buurten bij die zelfs 10 keer groter zijn dan het gemiddelde.



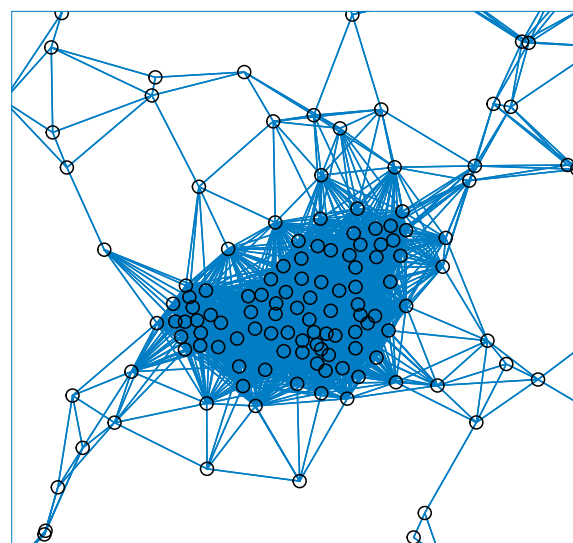
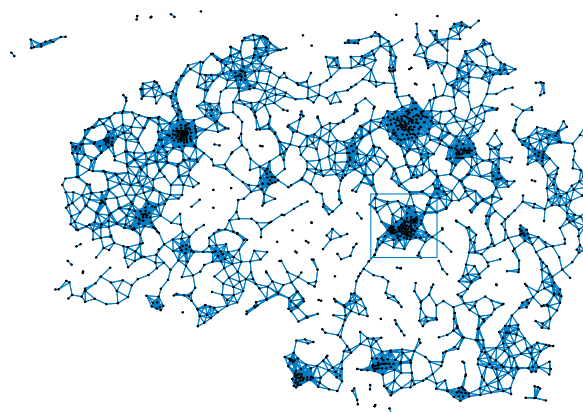
Figuur B4.5 Oppervlakte per buurt in km².

Om de ruimtelijke autocorrelatie te beoordelen, moest nader worden gedefinieerd wat werd bedoeld met twee waarnemingen die dicht bij elkaar liggen. Hiervoor is een afstandsmaat van 0 tot 1 en van 0 tot 3 kilometer getest. Deze werd gebaseerd op de verkenning van: de data in Qgis waarin sprake was clustering van verschillende variabelen; de correlogrammen (figuur. B.4.1 tot en met B.4.4); de resultaten van de Global Moran's Tests (paragraaf B.4.4.2).

De afstandsintervallen zijn verwerkt in een gewichtsmatrix die de relaties definieerde tussen de zwaartepunten (*centroids*) van de geometrische buurten. Een aanzienlijk probleem dat optrad bij de keuze voor een gewichtenmatrix (*spatial weight matrix*) – en dus bij de conceptualisatie van de ruimtelijke relaties – is dat sommige buurten dusdanig groot waren dat een te kleine afstandsmaat van bijvoorbeeld 1 kilometer vanuit het geometrische zwaartepunt niet buiten de eigen buurtgrenzen reikt. Dit heeft als gevolg dat te grote buurten



Figuur B4.6 Buurtcentroids in het onderzoeksgebied met buurt-koppelingen op basis van een afstandsinterval van 0-1km. Boven een plot van het gehele onderzoeksgebied met een uitsnede die onder is uitvergroot.



Figuur B4.7 Buurtcentroids in het onderzoeksgebied met buurt-koppelingen op basis van een afstandsinterval van 0-3km. Boven een plot van het gehele onderzoeksgebied met een uitsnede die onder is uitvergroot.

geen koppeling maken met een aangrenzende buurt terwijl dit wel gewenst is ten behoeve van een ruimtelijke analyse die de realiteit weerspiegelt (geografische buurten op het vasteland zijn nooit volledig geïsoleerd).

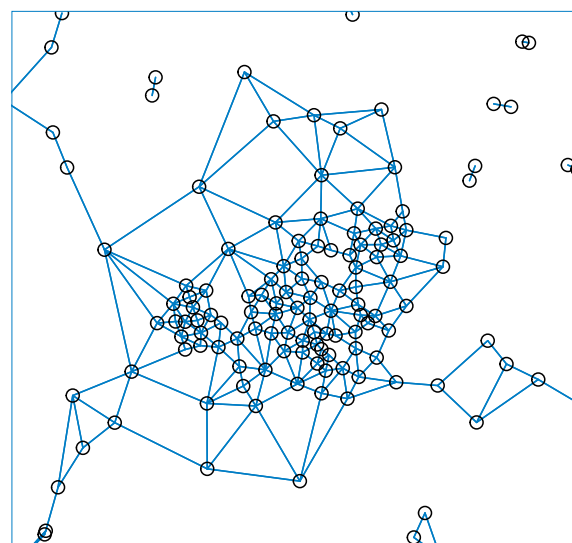
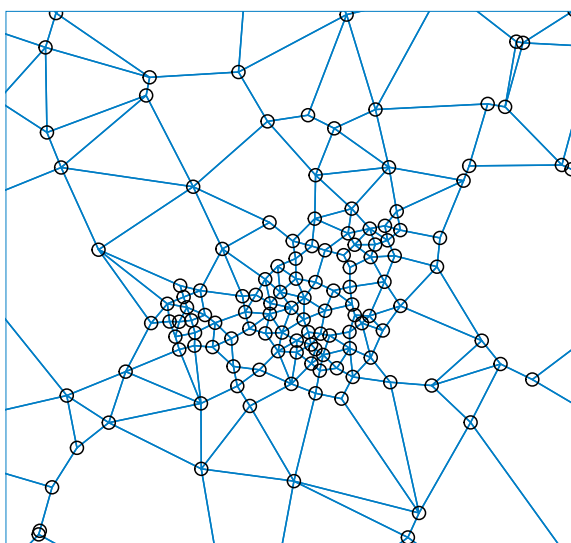
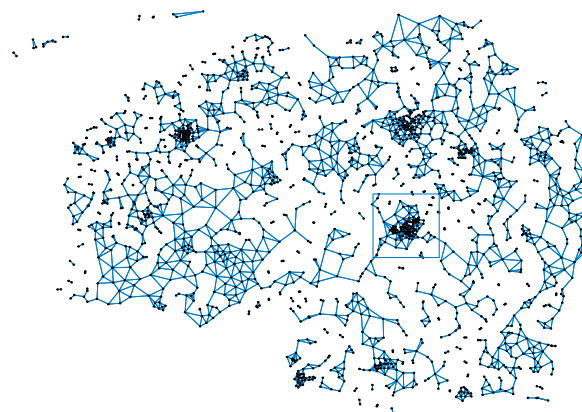
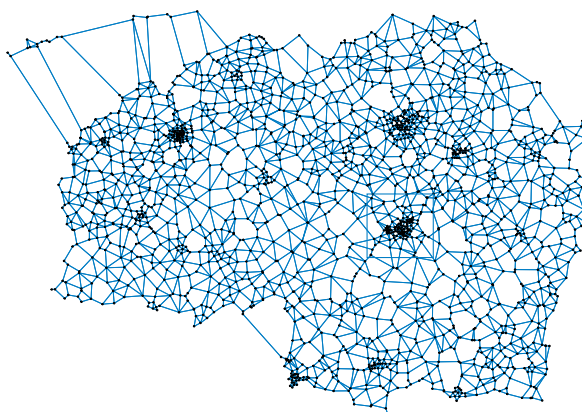
Deze problematiek blijkt ook uit de resultaten (figuur B.4.6 en B.4.7) van de gehanteerde afstandsmaten. Bij het afstandsinterval '0-1km' werden bijna negenhonderd buurten niet gekoppeld en gemiddeld een tot twee koppelingen per buurt gemaakt. Bij het afstandsinterval '0-3km' waren er nog maar dertien buurten zonder koppeling maar wel ruim elf koppelingen gemiddeld per buurt met een maximum van tweeënzeventig koppelingen.

Zowel ontbrekende koppelingen als overmatige koppelingen zorgen voor een suboptimale conceptualisatie van de ruimtelijke relaties op buurtniveau. Het probleem van ongewenste aantallen koppelingen is een gevolg van heterogene geometrische vormen van de buurten. Een mogelijke oplossing voor dit probleem was een slimmere ruimtelijke

gewichtenmatrix zoals de *Gabriel* matrix (Matula & Sokal, 1980), *Sphere of Influence* (Toussaint, 2014) of de *Relative Neighbour* matrix (Toussaint, 1980).

De zojuist geïntroduceerde 'slimme' conceptualisaties houden op basis van geometrische berekeningen gebaseerd op de zogenaamde *Delaunay*-triangulatie¹ rekening met buurtcentroids die op onregelmatige afstand van elkaar te vinden zijn (Ardilly, et al., 2018). Bij de *Gabriel*-methode komt dit neer op het vinden van een dichtsbijzinde buurt binnen de *Delaunay*-triangulatie. Er worden dan koppelingen gemaakt met dichtsbijzinde buurten in verschillende richtingen (grofgezegd langs de kortste driehoeksijden). In figuur B.4.8 is de toepassing van de *Gabriel* matrix op het onderzoeksgebied weergegeven. Het aantal gevormde koppelingen per buurt met andere buurten varieert van een tot en met acht met een gemiddelde aantal koppelingen per buurt van vier.

¹ In geometrie is een triangulatie een onderverdeling van een vlak object in driehoeken.



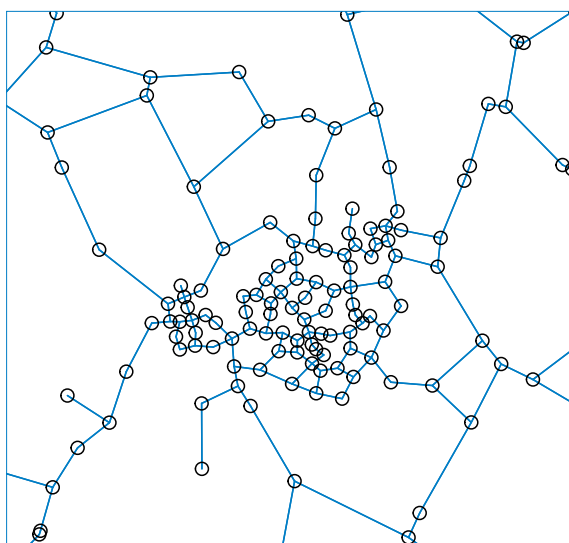
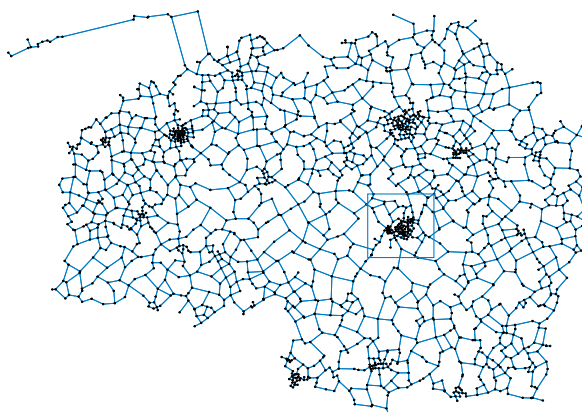
Figuur B4.8 Buurtcentroids in het onderzoeksgebied met buurt-koppelingen op basis van een *Gabriel* gewichtenmatrix voor burens. Boven een plot van het gehele onderzoeksgebied met een uitsnede die onder is uitvergroot.

Figuur B4.9 Buurtcentroids in het onderzoeksgebied met buurt-koppelingen op basis van een *Sphere of Influence* gewichtenmatrix voor burens. Boven een plot van het gehele onderzoeksgebied met een uitsnede die onder is uitvergroot.

De *Sphere of Influence* geometrie koppelt twee buurtcentroids als hun 'cirkels op basis van de dichtstbijzijnde buurt' elkaar overlappen (Ardilly, et al., 2018). In figuur B.4.9 is de toepassing van de *Sphere of Influence* matrix op het onderzoeksgebied weergegeven. Het aantal gevormde koppelingen per buurt met andere buurten varieert van een tot en met negen met een gemiddelde aantal koppelingen per buurt van ruim drie. Nadeel van deze conceptualisatie is dat driehonderdvijf buurten met slechts 1 koppeling relatief geïsoleerd blijven terwijl dit in werkelijkheid niet zo is.

De benadering van de *Relative Neighbour* gewichtenmatrix (figuur B.4.10) lijkt meer in de buurt te komen van de daadwerkelijke ruimtelijke relaties met een gemiddelde aantal koppelingen per buurt tussen de twee en drie. De berekening van burens is gebaseerd op het voldoen aan de formule:

$$d(p_i, p_j) \leq \max[d(p_i, p_k), d(p_j, p_k)] \quad \forall k = 1, \dots, n \quad k \neq i, j$$



Figuur B4.10 Buurtcentroids in het onderzoeksgebied met buurt-koppelingen op basis van een *Relative neighbour* gewichtenmatrix voor burens. Boven een plot van het gehele onderzoeksgebied met een uitsnede die onder is uitvergroot.

Met $d(p_i, p_j)$ als de afstand tussen de buurten p_i en p_j . Ondanks de verbetering, zijn er nog vierenzestig buurten met slechts één koppeling. Dit valt echter mee ten opzichte van de vorige conceptualisaties. Toussaint (1980) legt uit dat de *Relative Neighbour* gewichtenmatrix zich beter aanpast aan data door het kleinste aantal koppelingen te vereisen.

De als laatst besproken conceptualisaties gebaseerd op 'slimme' geometrische berekeningen voorkomen niet alleen een overschot maar ook een tekort aan koppelingen, ongeacht de onregelmatigheid van de afstanden tussen de buurtgebieden. Daarmee leek de keuze voor een 'slimme' ruimtelijke benadering logischer of gewenster dan de conceptualisaties gebaseerd op vaste afstanden. De methoden op basis van de vastgestelde afstandsmaten in figuur B.4.6 en B.4.7 zijn echter in gebruik gebleven, omdat een aantal praktische problemen optraden in een poging te corrigeren voor ruimtelijke autocorrelatie in poissonregressies. Dit zal later worden besproken, nadat eerst de ruimtelijke autocorrelatie voor de verschillende conceptualisering van ruimtelijke relaties is toegelicht.

B.4.4.2 Global Moran's Toets

Een Moran's I coëfficiënt groter dan de toetsing-grootheid $-1/(n-1)$ wijst op positieve ruimtelijke autocorrelatie, een Moran's I kleiner dan $-1/(n-1)$ wijst op negatieve ruimtelijke autocorrelatie.

Als toetsingsgrootheid voor de regio Noord-Nederland met 1994 buurten in de ruimtelijke analyse geldt:

$$-1/(n-1) = -5,0176e-4$$

Met behulp van gewichtenmatrixen (Tiefelsdorf, Griffith, & Boots, 1999; Kelejian & Prucha, 2010) kon de Moran I toets onder randomisatie worden uitgevoerd met het spDep softwarepakket voor 'R' (Bivand & Piras, 2015; Bivand & Wong, 2018). De berekeningen zijn in eerste instantie uitgevoerd voor alle *query*-resultaten en de Top25-*query*-resultaten als afhankelijke variabele van een poisson model. Ter aanvulling zijn ook de tien resultaten van de deelthema's meegenomen. Twee plus tien poissonmodellen maal de vijf ruimtelijke conceptualisaties geeft zestig verschillende testresultaten die zijn samengevat in tabel B.4.1. De aandacht ging vooral uit naar het totale aantal Top25 resultaten. Aangezien hier de meeste innovatieve startups werden verwacht.

Tabel B.4.1 Global Moran toetsing naar verschillende ruimtelijke conceptualisaties.

Gewichten-matrix: Poisson Model*	afstandsmaat '0-1km'		afstandsmaat '0-3km'		<i>Gabriel</i>		<i>Sphere of Influence</i>		<i>Relative Neigh</i>	
	Moran's I	p	Moran's I	p	Moran's I	p	Moran's I	p	Moran's I	p
Totaal (34%)	4,91e-2	0,07	7,13e-3	0,29	2,24e-2	0,07	2,39e-2	0,11	2,36e-2	0,11
Top25 (17%)	-2,65e-2	0,76	-3,04e-2	0,98	2,76e-3	0,42	-1,91e-2	0,82	1,05e-2	0,28
Top25 per TS										
Agri F (4%)	1,26e-2	0,34	2,17e-2	0,93	1,22e-2	0,76	7,46e-2	0,63	-1,95e-2	0,83
Chem (1%)	1,62e-2	0,30	2,67e-2	0,02	2,60e-2	0,04	2,72e-2	0,07	2,23e-2	0,11
C Indu (7%)	1,32e-2	0,63	9,65e-3	0,23	4,15e-3	0,38	7,21e-3	0,34	1,01e-2	0,29
Ener (5%)	3,21e-2	0,17	7,91e-3	0,27	1,37e-2	0,18	1,10e-2	0,28	3,57e-2	0,03
HTSM (4%)	1,07e-2	0,61	3,11e-2	0,01	2,55e-2	0,05	4,42e-2	0,01	5,96e-2	<0,01
ICT (5%)	4,31e-3	0,54	6,27e-3	0,31	3,24e-2	0,02	3,21e-2	0,05	4,07e-2	0,02
LS&H (4%)	4,92e-2	0,07	4,02e-2	<0,01	9,21e-4	0,46	2,37e-2	0,10	1,16e-2	0,27
Logi (2%)	2,11e-2	0,71	4,29e-3	0,34	2,16e-2	0,08	3,01e-2	0,06	3,56e-2	0,03
Tuinb (1%)	1,88e-2	0,27	1,65e-2	0,11	4,54e-3	0,37	6,46e-3	0,35	8,99e-3	0,31
Water (3%)	3,52e-3	0,45	-5,54e-3	0,64	-9,51e-3	0,71	-1,46e-2	0,76	-8,34e-3	0,65

* Met het percentage buurten dat minimaal 1 startup bevat ($\neq 0$) van de in totaal 1994 buurten, ter indicatie van eventuele onbetrouwbaarheid van de poisson modellen.

De uitsplitsingen naar deelthema gaven mogelijk minder betrouwbare resultaten vanwege het geringe aantal overgebleven resultaten. Deze worden in eerste instantie buiten beschouwing gelaten. Bij de eerste twee poisson modellen in de tabel m.b.t. de totale (en de Top25) resultaten was onder de alle ruimtelijke conceptualisaties sprake van een Moran I coëfficiënt groter dan de toetsingsgrootte wat wijst op positieve ruimtelijke autocorrelatie. Er was echter geen sprake van significantie. De afstandsmaat van '0-1km' en de Barel methode gaf echter wel de laagste p-waarde voor het totale model ($p=0,07$) maar nog steeds niet significant.

Het lijkt erop dat de *Gabriel* en *Relative Neighbour* methoden over het algemeen de meest significante resultaten van positieve ruimtelijke autocorrelatie hebben opgeleverd.

B.4.4.3 Modelleren van de verstoringen

Lambert et al. (2010) onderkennen een tekort aan ruimtelijk econometrische benaderingen beschikbaar om ruimtelijk gecorreleerde verstoring in telmodellen (zoals voor dit onderzoek in de poissonmodellen) te modelleren. Echter bleek de voorgestelde methode met bijhorende software beperkte uitvoerbaar. Als alternatief is gekeken naar de methode zoals voorgesteld door Dormann et al. (2007) maar deze bleek alleen geschikt voor de gewichtenmatrixen met afstandsmaten.

Specifiek is 'autovariate' regressie toegepast², met de opmerking van de auteurs dat deze tot nu toe hun weg nog niet hebben gevonden naar de literatuur. De auteurs staan niet in voor de statistische correctheid van deze benadering, dit blijft onzeker. In R biedt de functie 'autocov_dist {spdep}' de mogelijkheid om een nieuwe verklarende variabele te maken die wordt ingevoerd naast de andere verklarende variabelen in het model. De resultaten voor de afstandsmaat 0-3km zijn vergeleken in tabel B.4.2 met poissonmodellen zonder autocovariabele.

Uit tabel B.4.2. blijkt dat de verschillen klein zijn. gezien het gebrek aan zekerheid over de nauwkeurigheid van de methode en de kleine kleinere aantallen bij de opsplitsingen per topsector is besloten af te zien om eventuele ruimtelijke verstoringen te modelleren.

² De instructies waren oorspronkelijk bedoeld voor ecologen waarvan er veel geen statistici zijn.

Tabel B.4.2 Vergelijking tussen poissonmodellen met en zonder autocovariabele.

Voorspellers	ALLES BU WEBUNIEK			ALLES BU WEBUNIEK			ALLES BU Top 25 WEBUNIEK			ALLES BU Top 25 WEBUNIEK		
	Schatting	SE	p	Schatting	SE	p	Schatting	SE	p	Schatting	SE	p
(Intercept)	-0,68	0,09	<0,001	-0,55	0,08	<0,001	-1,08	0,11	<0,001	-1,14	0,11	<0,001
Bedrijfsinschrijvingen 2016(x1.000)	6,18	0,16	<0,001	6,24	0,16	<0,001	6,07	0,21	<0,001	6,13	0,21	<0,001
Omgevings-adresdichtheid (x100.000)	11,86	4,7	0,012	13,13	4,66	0,005	6,35	5,68	0,263	4,59	5,71	0,421
Aantal hoogopgeleiden (x100.000 in gemeente)	-1,14	0,4	0,005	-1,18	0,4	0,003	-0,67	0,49	0,17	-0,65	0,5	0,196
Aantal cafés binnen 1 km(x 10)	-0,28	0,03	<0,001	-0,28	0,03	<0,001	-0,17	0,03	<0,001	-0,2	0,03	<0,001
(1,2]	-0,32	0,06	<0,001	-0,33	0,06	<0,001	-0,31	0,08	<0,001	-0,3	0,08	<0,001
(2,3]	-0,4	0,11	<0,001	-0,43	0,11	<0,001	-0,5	0,15	0,001	-0,48	0,15	0,001
(3,4]	-0,62	0,28	0,027	-0,64	0,28	0,023	-1,43	0,59	0,015	-1,42	0,59	0,016
(1,2]	-0,21	0,08	0,009	-0,27	0,08	<0,001	-0,21	0,1	0,035	-0,19	0,1	0,064
(2,3]	-0,21	0,08	0,011	-0,28	0,08	<0,001	-0,38	0,11	<0,001	-0,34	0,1	0,001
(3,4]	-0,34	0,13	0,011	-0,43	0,13	0,001	-0,84	0,2	<0,001	-0,79	0,2	<0,001
Aantal westerse allochtonen(x 1.000)	-2,51	0,22	<0,001	-2,64	0,21	<0,001	-3,28	0,28	<0,001	-3,24	0,28	<0,001
Aantal westerse allochtonen(x 10.000 in wijk)	2,7	0,52	<0,001	3,1	0,51	<0,001	4,71	0,63	<0,001	4,62	0,63	<0,001
Groningen KI	1,03	0,23	<0,001	1,13	0,22	<0,001	1,22	0,27	<0,001	1,1	0,28	<0,001
Leeuwarden KI	0,74	0,15	<0,001	0,69	0,14	<0,001	1,14	0,17	<0,001	1,13	0,17	<0,001
autocovariabele	0,1	0,03	<0,001	Zonder autocovariabele			-	-	-	Zonder autocovariabele		
autocovariabele T25	-	-	-				-0,14	0,04	<0,001			
Observaties / ≠ 0	1994 / 685			1994 / 685			1994 / 342			1994 / 342		
Cox & Snell's R2 / Nagelkerke's R2	0,616 / 0,690			0,614 / 0,687			0,475 / 0,551			0,471 / 0,546		
Deviance	2549,498			2562,329			2656,173			2672,078		
AIC	4280,605			4291,436			3597,677			3611,583		

