



Limperg Instituut

Limperg Symposium Statistical Auditing 2019

“Veel afwijkingen verwacht – wat nu?”

22 mei 2019

Vrije Universiteit Amsterdam

Auditorium

De Boelelaan 1105

1081 HV Amsterdam

Programmaboekje



Inleiding

Tijdens de voorbije elf symposia was steeds een terugkerend thema, de vraag of en hoe controlewerkzaamheden kunnen worden gereduceerd als er subjectieve gunstige voorkennis aanwezig is over de kwaliteit van de te controleren populatie. Het audit risk model, en Bayesiaanse alternatieven, zijn er allemaal op gebaseerd gunstige voorkennis om te zetten in efficiëntere controle.

Voorkennis

Maar wat nu, als die gunstige voorkennis afwezig is, of, sterker nog, als eigenlijk wel duidelijk is dat de te controleren gegevens fouten bevatten? Steekproeven zijn eigenlijk alleen maar efficiënt als goedkeuren bij nul fouten in de lijn der verwachting ligt. Wanneer de toelaatbare afwijking gelijk wordt gesteld aan de uitvoeringsmaterialiteit, leidt het kleinste foutje immers al tot een maximale fout die boven de uitvoeringsmaterialiteit uitkomt.

Zelfs als de aanwezige fout niet materieel is, kan een steekproef die alleen bij 0 fouten tot goedkeuring leidt, al een teleurstellend resultaat opleveren. Neem bijvoorbeeld een uitvoeringsmaterialiteit van 3 procent bij 95 procent betrouwbaarheid: daarvoor is de kleinst mogelijke steekproefomvang 100 bij 0 toegestane fouten in de steekproef. Als de werkelijke fout in de populatie 1 procent is, is het verwachte aantal fouten in de steekproef 1. De uitkomsten 0 en 1 zijn even waarschijnlijk en beide zijn de meest voorkomende met 37 procent kans. Dus 63 procent van alle steekproeven uit een populatie met 1 procent fouten zal niet leiden tot goedkeuring als de uitvoeringsmaterialiteit 3 procent is bij 95 procent betrouwbaarheid!

Het wordt nog veel erger als de werkelijke fout boven de uitvoeringsmaterialiteit ligt. Als de populatie 5 procent fouten bevat, is de kans op 0 fouten in 100 waarnemingen kleiner dan 1 procent. Gelukkig maar: bij 3 procent is die kans immers 5 procent.

Data-analyse

De eerste vraag die tijdens het symposium zal worden voorgelegd is of steekproeven überhaupt nog nut hebben in dit soort situaties, waarbij met materiële fouten rekening moet worden gehouden. Vervolgens wordt ingegaan op alternatieven die wij de verzamelaar data-analyse geven.

Met data-analyse kan men integraal fouten uit een bestand halen, wordt vaak gesteld. En dat is ook zo, maar dat geldt alleen voor die situaties waarin bekend is hoe een fout eruit ziet. Welke karakteristieken hebben foute transacties met elkaar gemeen en onderscheiden foute transacties van correcte? Hoe voorkomen we bij het apart zetten van fouten dat er correcte transacties meegaan? En, hoe weten we dat we alle fouten apart hebben gezet?

Tijdens de drie casestudies over data-analyse in situaties waarin veel fouten worden verwacht, worden deze vragen specifiek gemaakt en beantwoord. Uitdaging is om ons daarbij eerst te realiseren dat data-analyse maatwerk is, maar pas efficiënt wordt als er regels uit zijn af te leiden die te generaliseren zijn.

Essaybundel

De Stuurgroep publiceert ongeveer maandelijks een column op de Accountant, <https://www.accountant.nl/themas/statistical-auditing/>. Er is een bundel gemaakt van deze columns en van andere artikelen van de leden van de Stuurgroep. Deze essaybundel, epub: https://limperginstituut.nl/?media_dl=980, pdf: https://limperginstituut.nl/?media_dl=1001, zal worden gepresenteerd op dit symposium.

[De tekst van deze inleiding is gebaseerd op column 74, <https://www.accountant.nl/artikelen/2019/4/euwe-pagina--symposium-statistical-auditing-2019-als-het-allem/>, van Stuurgroep Statistical Auditing gebruikt. De auteur van column 74 is Paul van Batenburg, werkzaam bij Deloitte en lid van de Stuurgroep.]

Inhoud

Inleiding	2
Voorkennis	2
Data-analyse	2
Essaybundel	3
Inhoud	4
Essaybundel Statistical Auditing	5
De stuurgroep Statistical Auditing maakt de balans op na 25 jaar en bundelt eigen werk van de leden in een essaybundel	5
Wiskundige steekproeven	5
Columns	5
E-boek	6
Programma	7
Dagvoorzitter en sprekers	8
Samenvatting van de presentaties	10
Presentaties	
Steekproeven als veel fouten worden verwacht, Paul van Batenburg	13
Brute rekenkracht voor de audit van tussenrekeningen, Jacques de Swart	26
Met Bayesiaanse statistiek het kaf van het koren scheiden, Eric-Jan Wagenmakers	38
Zelflerende algoritmen bij het verbeteren van datakwaliteit, Willem Jan Willemse	50
Kunstmatige Intelligentie ontrafeld, Tjerk de Greef	62

Essaybundel Statistical Auditing

De stuurgroep Statistical Auditing maakt de balans op na 25 jaar en bundelt eigen werk van de leden in een essaybundel

Na afloop van de presentatie van de Green Paper over auditkwaliteit door de NBA hebben leden van de stuurgroep Statistical Auditing aangeboden om te verzamelen wat er vanuit dat vakgebied - dat houdt in: data-analyse, wiskunde en statistiek - aan auditkwaliteit al is en nog kan worden bijgedragen. Die vraag leidde tot introspectie bij de stuurgroep: wat hebben de leden van de stuurgroep in de afgelopen 25 jaar gepubliceerd dat aan deze vraag raakt?

Zie hier het antwoord. De stuurgroep heeft een redactie gevormd en die heeft eigen artikelen (waarvan een of meer auteurs lid zijn of waren van de stuurgroep) en columns op <https://www.accountant.nl/themas/statistical-auditing/> gebundeld. Met een inleiding en enkele aanvullende artikelen is dit de Essaybundel Statistical Auditing geworden. In de bundel wordt ook naar reeds verschenen literatuur verwezen die door anderen is gepubliceerd.

De inleiding van de Essaybundel Statistical Auditing behandelt de hoofdlijnen van het controleproces aan de hand van een netwerkmodel. Met behulp van dat netwerkmodel beschrijven de auteurs de beoordeling van de interne beheersing, inclusief de IT, de gegevensgerichte detailcontroles en de gegevensgerichte cijferanalyse. Daarnaast wordt ingegaan op de reductie van de hoeveelheid detailcontroles door middel van risicoanalyse (Audit Risk Model dan wel Bayesiaans). Het inleidende artikel sluit af met een suggestie hoe en waar door middel van (wetenschappelijk) onderzoek de praktijk kan leren van de theorie en andersom.

Wiskundige steekproeven

Het artikel over fundamenteel onderzoek naar wiskundige steekproeven focust op de theoretische kant van de wiskunde met betrekking tot geldsteekproeven. Het verwijst naar de complexe wiskunde die een aantal mogelijke varianten van het extrapoleren van de bovengrenzen verklaart, en verwijst belangstellenden naar het desbetreffende onderzoek zonder de formules en wiskundige bewijzen weer te geven die noodzakelijkerwijs bij deze materie horen.

Ook bevat de bundel drie artikelen die eerder in het MAB zijn verschenen en een artikel uit Handboek Accountancy. De MAB-artikelen gaan over risicoanalyse in de accountantscontrole, statistische technieken voor het bijsturen van de interne beheersing en de effecten van een goede interne beheersing op volgende controles. Het artikel uit Handboek Accountancy gaat over de relatie van het begrip materialiteit met de operationele grootheden uitvoeringsmaterialiteit en toelaatbare afwijking.

Columns

Een belangrijk deel van het gedachtegoed van de Stuurgroep is vastgelegd in columns. De redacteuren van de bundel hebben er veel van in de bundel opgenomen. De meeste daarvan gaan over gegevensgerichte detailcontroles en cijferanalyse (of zo men wil data-analyse), kortom, controletechnieken die gerelateerd zijn aan kwantitatieve methoden en technieken.

In de bundel heeft de redactie zich alleen bekommerd om financiële controles. Zoals ook uit de columns blijkt kan dat bijvoorbeeld een jaarrekeningcontrole zijn, een fiscale controle en uiteraard ook een fraudeonderzoek.

E-boek

De bundel is gepubliceerd als e-boek. Er is een beperkt aantal papieren exemplaren gedrukt. Het boek wordt tijdens het symposium aan de directeur van de Foundation for Auditing Research prof. dr. Jan Bouwens en aan de voorzitter van het Limperg Instituut prof. dr. Tom Groot.

De essaybundel is elektronisch beschikbaar op de site van het Limperg Instituut, epub: https://limperginstituut.nl/?media_dl=980, pdf: https://limperginstituut.nl/?media_dl=1001.

[Voor dit hoofdstuk over de essaybundel is de tekst van column 75, <https://www.accountant.nl/artikelen/2019/5/essaybundel-stuurgroep-statistical-auditing/>, gebruikt. Auteur van column 75 is Hein Kloosterman.]

Hein Kloosterman RE RA heeft een verleden als vaktechnisch adviseur controle, IT-audit en Statistical Audit bij de Belastingdienst. Hij is betrokken bij de accountantsopleiding van de Business Universiteit Nyenrode en de Erasmus Universiteit (ESAA) en bij de IT-auditopleiding van de VU. Verder is hij lid van de redactieraad van Handboek EDP-audit en van de Stuurgroep Statistical Audit. Hij werkt als zelfstandig adviseur op het gebied van IT-audit en Statistical Audit; hij adviseert daarbij ook aan en namens de KEY Group.]

Programma

Plaats: Auditorium, Vrije Universiteit Amsterdam,
De Boelelaan 1105,
1081 HV, Amsterdam.

Datum: 22 mei 2019

Tijd: 13:30 tot 17:20, met ontvangst vanaf 13:00 en borrel tot 18:00

13:00 Ontvangst met koffie en thee

13:30 Opening,
Tom Groot, bestuursvoorzitter Limperg Instituut

13:35 Inleiding,
Jan Bouwens, dagvoorzitter, UvA

13:45 *Steekproeven als veel fouten worden verwacht:
met een dobbelsteen door een mijnenveld lopen?*
Paul van Batenburg, Deloitte

14:20 *Brute rekenkracht voor de audit van tussenrekeningen,*
Jacques de Swart, PwC

14:55 Pauze

15:25 *Met Bayesiaanse statistiek het kaf van het koren scheiden,*
Eric-Jan Wagenmakers, UvA

16:00 *Zelflerende algoritmen bij het verbeteren van datakwaliteit,*
Willem Jan Willemse, DNB

16:35 *Kunstmatige Intelligentie ontrafeld,*
Tjerk de Greef, Legal Intelligence/WoltersKluwer

17:10 Afsluiting,
Jan Bouwens

17:20 Drankje en hapje

18:00 Einde borrel

Dagvoorzitter en sprekers

Prof. dr. J.F.M.G. (Jan) Bouwens is professor of accounting aan de Amsterdam Business School en research fellow aan de Judge Business School van The University of Cambridge waar hij eveneens doceert en onderzoek uitvoert. Tijdens het academisch jaar van 2013/2014 was hij visiting professor aan de Harvard Business School. Van 2004 tot 2014 was hij Professor of Financial and Management Accounting aan Universiteit van Tilburg. Tussen 1999 en 2004 werkte hij als hoogleraar Accounting aan Nyenrode, waar hij tevens de onderzoeksgroep leidde. Hij schreef zijn proefschrift in managerial accounting aan de Universiteit van Tilburg. Hij heeft een tiental publicaties in accountingtijdschriften van wereldfaam staan, en schrijft ook voor een breder publiek in kranten en vaktijdschriften. Jan is thans bestuurslid van de Foundation for Auditing Research. Dit instituut legt zich toe op vergroting van de kennis omtrent auditing.

Paul van Batenburg is statisticus bij Deloitte. Hij adviseert accountants, fiscalisten en consultants bij het gebruik van statistiek. Paul is een van de auteurs van de NBA- Handreiking 1141 Data-analyse bij de controle. Door een samenwerking met V&A accountants-adviseurs werkt Paul ook voor Tier 2 en 3 accountantskantoren bij het opzetten, uitvoeren en evalueren van gegevensgerichte controles. Hij geeft gastcolleges over dit onderwerp op vier universiteiten. Voor meer informatie zie www.steekproeven.eu.

Prof. dr. J.J.B. (Jacques) de Swart MBA De missie en passie van Jacques is het ontsluiten van de pracht en kracht van wiskunde op grootse en meeslepende wijze. Vanuit diverse rollen probeert hij deze missie te volbrengen: als voorzitter van de Stuurgroep Statistical Auditing die dit symposium organiseert, als hoogleraar Toegepaste Wiskunde aan Nyenrode Business Universiteit en als consultant binnen PwC NL waar hij verantwoordelijk is voor de Data Analytics groep binnen Consulting. Jacques heeft een achtergrond als programmeur en wiskundige.

Eric-Jan Wagenmakers is wiskundig psycholoog en toegewijd Bayesiaan. Hij is werkzaam als hoogleraar bij de afdeling Psychologische Methodenleer aan de Universiteit van Amsterdam, alwaar hij de ontwikkeling aanstuurt van JASP, een open-source software programma voor statistische analyses.

Dr. W.J. (Willem Jan) Willemse is dataspecialist binnen de divisie Toezicht Verzekeraars van De Nederlandsche Bank. Hij houdt zich bezig met het ontwikkelen en toepassen van nieuwe methoden en technieken in het toezicht. Hij is gepromoveerd in de Technische Informatica in Delft en heeft Actuariële Wetenschappen in Amsterdam gestudeerd.

Tjerk de Greef In zijn huidige rol als Director of Search Technology is Tjerk verantwoordelijk voor de toepassing van zoektechnologie binnen de wereldwijde Wolters Kluwer organisatie. In deze functie is hij geïnteresseerd in het toepassen van information retrieval innovaties voor alle divisies binnen Wolters Kluwer. Daarbij leunen veel innovaties op kunstmatige intelligentie. Na het behalen van zijn PhD werkte Tjerk als Chief Information Officer bij Legal Intelligence en trad hij op als Program Manager & Scientific Secretary bij het Nederlandse ICT-wetenschapsprogramma COMMIT/. Als onderzoeker wijdde hij zijn tijd aan Computer Science-onderzoek op het gebied van Human Factors & Human Computer Interaction.

Samenvatting van de presentaties

Paul van Batenburg: Steekproeven als veel fouten worden verwacht: met een dobbelsteen door een mijnenveld lopen?

Paul van Batenburg zal laten zien dat er ook bij populaties met hoge foutdichtheden nog best met steekproeven gewerkt kan worden. Als goedkeuren zonder correcties er niet in zit, kan de controleur wel een betrouwbare en voldoende nauwkeurige schatting maken van de noodzakelijke correctie om daarna goed te keuren. Dat kost veel werk, maar het onthouden van goedkeuring kost ook geld. Misschien is het wel het beste, zal ik laten zien, om gewoon maar eerst op 0 fouten te rekenen en dan de steekpref uit te breiden. In het tweede deel van mijn presentatie zal ik de AOQL methode laten zien. Die kost aanzienlijk meer werk, maar daar krijg je ook meer voor: een kwaliteitsgarantie. Alhoewel?

Jacques de Swart: Brute rekenkracht voor de audit van tussenrekeningen

Wie kent het niet? Bij de jaarlijkse sluiting van de boeken blijken er nog journaalposten in tussenrekeningen uitgezocht te moeten worden.

Een hels karwei dat nog maar al te vaak te lijf wordt gegaan met een spreadsheet. Sorteren op bijvoorbeeld factuurnummers en dan maar hopen dat er wat weggestreept kan worden. Niet zelden is dit zoekwerk bij de jaarafsluiting nog niet volledig afgerond en wordt de onopgeloste puzzel meegenomen het nieuwe boekjaar in. Wiskundig gezien is het niet vreemd dat het uitpluizen van tussenrekeningen eigenlijk onbegonnen werk is. Het gaat hier namelijk om een Mixed Integer Programming (MIP) probleem en zo'n probleem is NP-hard, wat zoveel betekent als dat de oplostijd harder dan polynomiaal stijgt als het aantal uit te zoeken journaalposten toeneemt.

In deze lezing behandel ik een aanpak inclusief voorbeeld om dit probleem met brute rekenkracht te lijf te gaan. Ook schets ik de ontwikkelingen die er voor zorgden dat deze aanpak nu praktisch toepasbaar is.

Eric-Jan Wagenmakers: Met Bayesiaanse statistiek het kaf van het koren scheiden

De Bayesiaanse benadering verleent de accountant een aantal unieke voordelen ten opzichte van de standaard "frequentistische" benadering.

Zo kan de Bayesiaanse accountant indien gewenst de priorverdeling aanpassen om verwachtingen omtrent het foutenpercentage in te bouwen in de analyse. Daarnaast kan de Bayesiaanse accountant ook blijven controleren totdat een gewenst niveau van zekerheid is bereikt. De planningsfase maakt derhalve geen noodzakelijk deel uit van de evaluatiefase. Tenslotte kan de Bayesiaanse accountant de mate van statistisch bewijs kwantificeren zonder te hoeven verwijzen naar hypothetische uitkomsten. In zijn presentatie zal Wagenmakers eerst kort de basisprincipes van de Bayesiaanse benadering toelichten, alvorens in te gaan op de toepassing binnen de auditing. Hij zal ook het programma JASP demonstreren, en met name de module "JASP for Audit", die bedoeld is om het leven van de accountant (zowel Bayesiaans als frequentist) te vergemakkelijken.

Willem Jan Willemse: Zelflerende algoritmen bij het verbeteren van datakwaliteit

DNB gebruikt in haar werk als toezichthouder rapportages van onder toezicht staande ondernemingen. Bij het beoordelen van de plausibiliteit van deze rapportages is dit jaar begonnen met de toepassing van machine learning-technieken.

Willem Jan zal drie zelflerende algoritmen toelichten en laten zien hoe deze in de praktijk worden ingezet bij de plausibiliteitsbeoordelingen van toezichtrapportages:

- dimensiereductie en clustering met het t-sne algoritme;
- uitbijteranalyse aan de hand van een one-class support vector machine; en
- patroonherkenning aan de hand van associatieregels

Deze algoritmen zijn vormen van unsupervised machine learning en vereisen daarmee geen vooraf geclassificeerde data. Van deze algoritmen worden voor- en nadelen toegelicht.

Tot slot gaat Willem Jan in op de uitdagingen die het gebruik van open source software en de resultaten daarvan in (grote) organisaties met zich meebrengen. Zo is er onder meer een noodzaak voor een gestructureerd beheer van omgevingen en codeversies en zijn afspraken nodig om de reproduceerbaarheid van resultaten te waarborgen.

Tjerk de Greef: Kunstmatige Intelligentie ontrafeld

De lezing geeft een grondige maar begrijpbare inleiding in kunstmatige intelligentie en Big Data. De vijf fundamentele basistechnieken in kunstmatige zullen niet alleen begrijpend en theoretisch uitgelegd worden, maar ook aan de hand van diverse commerciële voorbeelden. De introductie in Artificial Intelligence en Big Data gaat dus veel verder dan mooie 'buzz words'.

Het doel van de 'workshop' is na denken wat kunstmatige intelligentie voor uw praktijk betekent en stimuleert kritisch na te denken over de toegevoegde waarde van kunstmatige intelligentie voor het probleem dat u aan het oplossen bent. Kunstmatige Intelligentie is namelijk niet een doel maar een middel.

De slides zullen worden gedistribueerd en deze bevatten verschillende verwijzingen naar online publicaties en korte verklarende video's. Na de presentatie zou het luisterend publiek voldoende begrip moeten hebben om kritische vragen te stellen aan leveranciers van oplossingen voor kunstmatige intelligentie en big data, zonder daarbij te worden weggeblazen door mooie termen.



Limperg Instituut

Veel fouten verwacht: met een dobbelsteen door een mijnenveld?

Paul van Batenburg



Casus

- Europese subsidiestroom
- Instantie keert uit en claimt vervolgens bij EU
- Goedkeurende verklaring nodig
- Protocol vereist (uitvoerings-)materialiteit 2% bij 90% betrouwbaarheid
- Rechtmatigheid op basis van inspectie ter plaatse
- Ervaring leert dat 4% onrechtmatig wordt uitgekeerd
- Geld terug vragen kost meer dan het oplevert
- Instantie kan (gedeeltelijk) afzien van claim



Vraag 1

- Heeft het trekken van een steekproef nut?
(ja/nee/geen mening)
- Achtergrond:
 - Steekproef volgens protocol 139
 - 2% (uitvoerings-)materialiteit, 90% betrouwbaarheid, 10% vermoedelijke fout
 - Daar in mag 0,274 foutfractie worden gevonden
 - Maar, als de werkelijke fout 4% is, verwachten we 5,56 fouten in die 139...



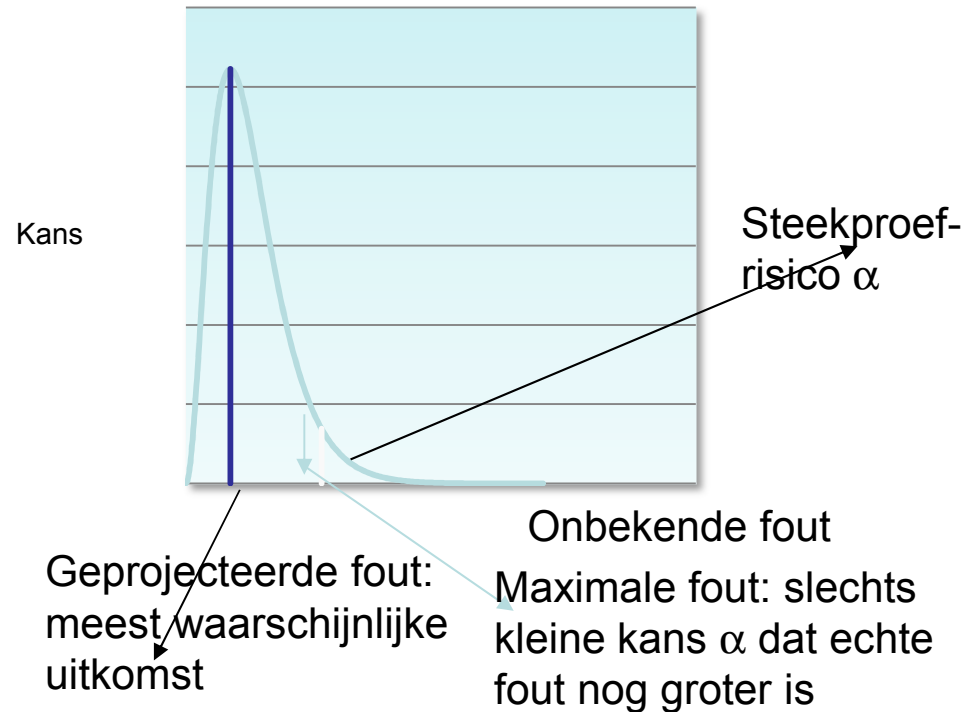
Limperg Instituut

Vier alternatieven

1. Grote(re) steekproef om noodzakelijke reductie in claim te schatten
2. Eerst kleine steekproef, dan later uitbreiden
3. AOQL methode
4. Googelen



De basis: het verkeerslicht



- **Groen:** maximale fout onder uitvoeringsmaterialiteit: goedkeuren en geprojecteerde fout ter correctie voorleggen
- **Oranje:** maximale fout minus geprojecteerde fout onder uitvoeringsmaterialiteit: goedkeuren na correctie van de geprojecteerde fout
- **Rood:** maximale fout minus geprojecteerde fout boven uitvoeringsmaterialiteit: werkzaamheden uitbreiden om in oranje te komen



Optie 1: Oranje

SMart Audit Sampling Helpfile SMASH		Audit inputs		2019 V1 December, 2018		AICPA Audit Sampling Guide (2017 ed.), app C	
Population amount	1.000.000	if B11>0 then D4 should equal D3					
Performance Materiality	60.000						
Tolerable Misstatement	60.000						
Sample design not allowing for overstatements				R factor	required assurance	sample size	interval
Levels of substantive testing	Standard	3,00	95%	50	20.000		
	Chosen	2,31	90%	39	25.974		
Sample design allowing for overstatements in population							
Presumed overstatement amount in population	Expected sample overstatement proportion	R factor	required assurance from sample (column M)	sample size	interval		
40.000	11,323	16,99	90%	284	3.531		

- Als de geprojecteerde fout op de claim in mindering wordt gebracht kan de toelaatbare afwijking 6% worden
- Zoek aantal fouten k en omvang n waarbij $k/n=4\%$ en $GAMMA.INV(90\%;k;1/n)=6\%$
- Dus = $IF(B11>0;(INDEX(M1:M17;I1+1)-I1*(INDEX(M1:M17;I1+2)-INDEX(M1:M17;I1+1)))/(D3/B11-(INDEX(M1:M17;I1+2)-INDEX(M1:M17;I1+1)));J1)$ uitrekenen
- Ruim 11 fouten in 284 waarnemingen geeft geprojecteerde fout 4% en de maximale fout 6% [Oranje]



Optie 2: Groen verwachten, Rood krijgen en dan uitbreiden voor Oranje

- Bij 2% (uitvoerings-)materialiteit, 90% betrouwbaarheid en 0 toegestane fouten volgt een minimale omvang van $2,31/0,02=116$
- Als populatie 4% fouten bevat, zal deze steekproef 4,64 fouten opleveren
- Geprojecteerde fout is dan 4% en de maximale fout ruim 7% [Rood]
- Steekproef uitbreiden met 167 geeft verwacht 6,68 extra fouten (totaal 11,32)
- Totale steekproef geeft geprojecteerde fout 4% en maximale fout 6% [Oranje]
- *En het aardige is dat $116+167+af ronding=284!$*

Evaluation at chosen confidence level		
Trial and error? Vuistregel 30% van n per fout	Final population amount	1.000.000
	Final sample size	116
	Interval	8.621
	Lower Error Limit	18.800
	Most Likely Error	40.000
	Upper Error Limit	76.255
	Work sufficient to accept population?	no
	Precision	36.255
	Work sufficient to estimate overstatement?	no
	Required entity adjustment	56.255
Sample expansion	167	
New Errors	6.680	
New Precision	19.992	



Voor alle zekerheid: taboe op uitbreiden?

- Uitbreiden van een steekproef om alsnog goed te keuren (Rood naar Groen) gaat ten koste van de gewenste betrouwbaarheid
- Dus niet hopen op een foutloze tweede steekproef om alsnog goed te keuren
- Ik breid de steekproef uit voor een betere schatting van de noodzakelijke correctie (Standaard 540, A94), en verwacht ook extra fouten (Rood naar Oranje)



Optie 3: AOQL

- Dodge & Romig (1933, 1957): formule voor de verwachte foutkans als populatie alleen bij 0 fouten in de steekproef wordt goedgekeurd en anders wordt opgeschoond tot er 0 fouten achter blijven
- Minimaal noodzakelijke steekproef om die verwachte foutkans aan norm $P\%$ te laten voldoen is $37/P$
- Om gerealiseerde foutkans daar dicht bij te krijgen zijn zeker 10 tot 20 deelpopulaties nodig
- Methode is altijd effectief, maar pas efficiënt als massa heterogeen is, en populatie wordt opgedeeld in kleine deelpopulaties met hoge en grote met lage foutkans

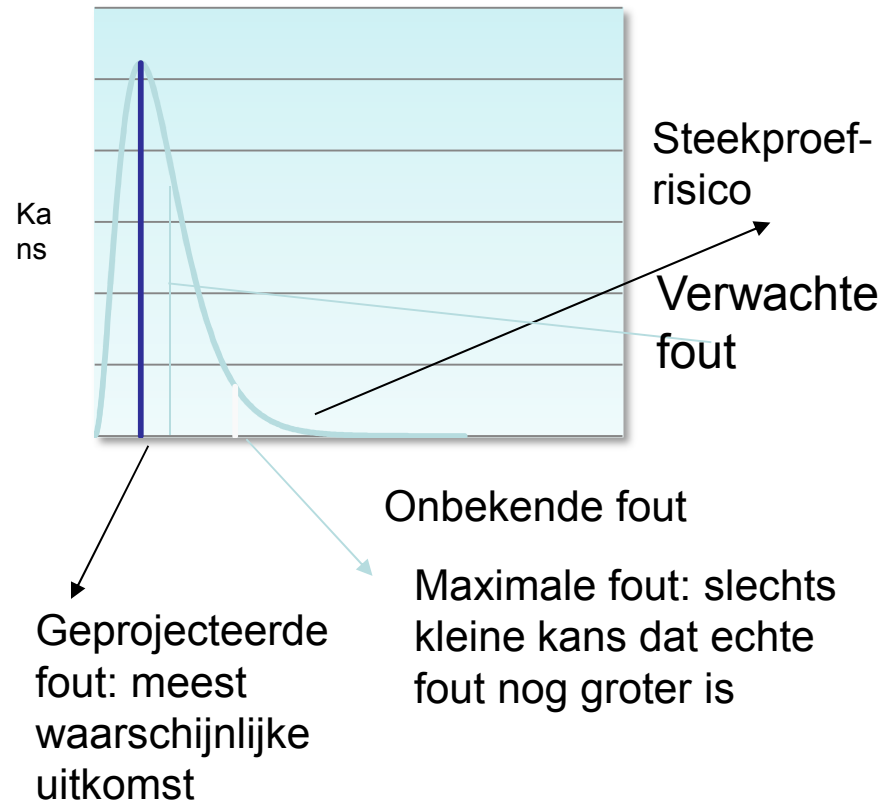


In dit voorbeeld

- Om te garanderen dat niet meer dan 2% fouten achterblijft is per deelpopulatie een steekproef van 19 nodig
- En moet elke deelpopulatie met fouten in de steekproef worden opgeschoond, desnoods door integrale controle
- Als de werkelijke foutkans 4% is, zal 47% van alle (10 a 20) steekproeven (van 19) worden goedgekeurd
- AOQL is dus alleen maar werkbaar als de foutkans per deelpopulatie varieert en geconcentreerd kan worden in kleine deelpopulaties met een hoge foutkans
- Minimale werklast 10 a 20 maal 19 plus opschonen van afgekeurde deelpopulaties
- Uitkomst is een zeker acceptabele populatie



Pas op voor appels en peren



- AOQL doet geen uitspraak over de geprojecteerde fout, maar over de verwachte (en gemiddelde) fout
- De Gamma(a,b) functie heeft als modus $(a-1)b$, verwachting ab , variantie ab^2 en als $100(1-\alpha)\%$ bovengrens $\text{GAMMA.INV}(1-\alpha; a; b)$



Vraag 2 en 3

- Heeft het trekken van een steekproef nut? (ja/nee/geen mening)
- Zo ja, welke optie heeft uw voorkeur? (1/2/3/nvt)
- *De screenshots komen uit spreadsheet SMASH dat te downloaden is van www.steekproeven.eu*
- *Zie ook columns 38 en 41 op Accountant.nl onder Statistical Auditing*



Toegift: methode 4

- Googelen op “steekproefomvang”
- Alle sites met werkende calculator die boven mijn url staan (www.steekproeven.eu)
- Waarom is dit minder dan 284? De aanname van normaliteit zorgt voor een dunnere rechterstaart

checkmarket.nl	2144
surveymonkey.nl	1576
peilingpraktijken.nl	370
steekproefcalculator.com	260
2refelct.nl	260
corpos.nl	257
allesovermarktonderzoek.nl	256
steekproeven.eu	260



Limpert Instituut

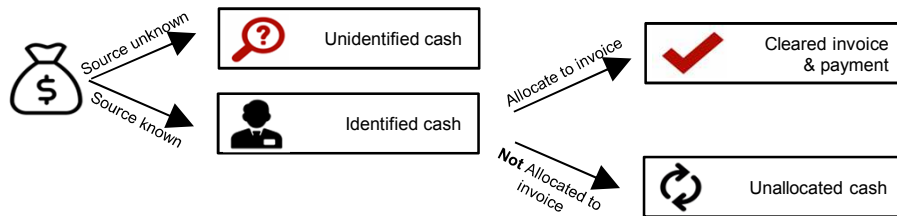
Brute rekenkracht voor de audit van tussenrekeningen

Jacques de Swart

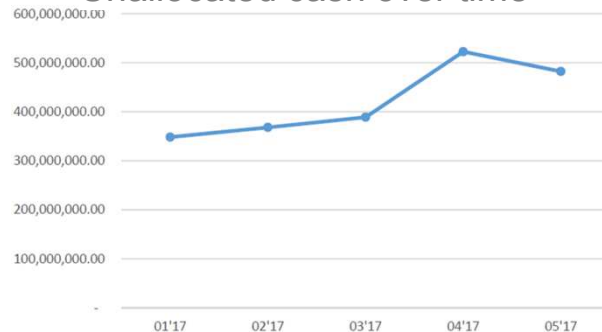


What was the problem?

An international client has an unallocated cash amount that is growing over time



Unallocated cash over time



So why is it a problem?

- Potential P&L impact (bad debt provisions/ gains)
- Risk of blocked sales orders: orders held when customer is over credit limit when payment not posted
- Potential wrong calculation of credit limits/ payment terms used
- Incorrect overdue reporting
- Possibility of incorrect dunning / collections management

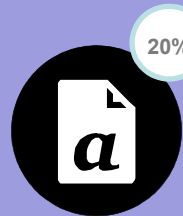
And.....is it an indication that the processes are not working as they are designed? (or not clearly defined....)



How did we solve the problem?

The Intelligent Matcher matches unallocated cash with outstanding invoices

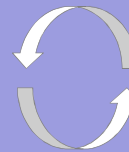
The Intelligent Matcher (TIM) matches debits and credits based on AI & Mixed Integer Programming.



Data is gathered and used as input for TIM. TIM does have a minimum amount of data fields and a standard set of rules required to ensure matches between the debit and credit amounts



Besides the basic matching rules, additional rules may be implemented into TIM to determine how matches can be made between debit and credit amount (amounts, dates, client names etc)



After running TIM matches will be presented based on the defined rules. Based on analysis, rules can be added to increase the amount of matches. Non-matched items are analyzed and explained in close cooperation with auditee



Limpert Instituut

TIM demo

www.pwc.nl/en/daapps#smart-processing



Did it work?

Next to root cause identification, the project resulted in restatement

Out of all unallocated cash for third parties, 94% of total value could be matched and/or explained

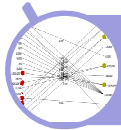
One OpCo revealed a positive P&L impact of 59% which lead to restatement

Analysis indicates conservative provisioning levels due to unallocated cash



Limperg Instituut

Why use TIM? Six reasons to work with TIM



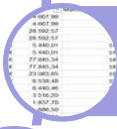
Computational efficiency drives high quality matches within limited time



Machine learning is applied in TIM to derive new matching logic inductively



Audit trail is captured by TIM



Clearing proposal is generated by TIM and can be uploaded for clearing



Extensive root cause analysis provides insight in to areas for process improvement



P&L impact is quantified



Why now (1/6)?

Now there is the ability of a machine to perceive its environment & perform tasks that normally require human intelligence

Sense



AI can see, hear, speak, smell, feel, understand gestures and even interface with your brain

Natural language

Audio and speech

Machine vision

Navigation

Visualization

Think



AI is helping us **make better decisions** – and doing it faster, better, more cheaply and more accurately

Knowledge and representation

Planning and Reasoning

Machine Learning

Deep Learning

Simulation and Digital Twins

Act



AI is **equaling or surpassing humans** in all sorts of tasks

- playing games, driving cars and making recommendations

Intelligent automation

Deep question and answering

Machine translation

Collaborative systems

Adaptive systems



Why now (2/6)?

Data analytics get used in enterprises beyond supporting evidence



Human in the Loop

Assisted Intelligence

AI systems that assist humans in making decisions or taking actions. Hard-wired systems that do not learn from their interactions.



No Human in the Loop

Automated Intelligence

Automation of manual and cognitive tasks that are routine. This does not involve new ways of doing things– automates existing tasks

**Hardwired /
Specific
Systems**

**Adaptive
Systems**

Augmented Intelligence

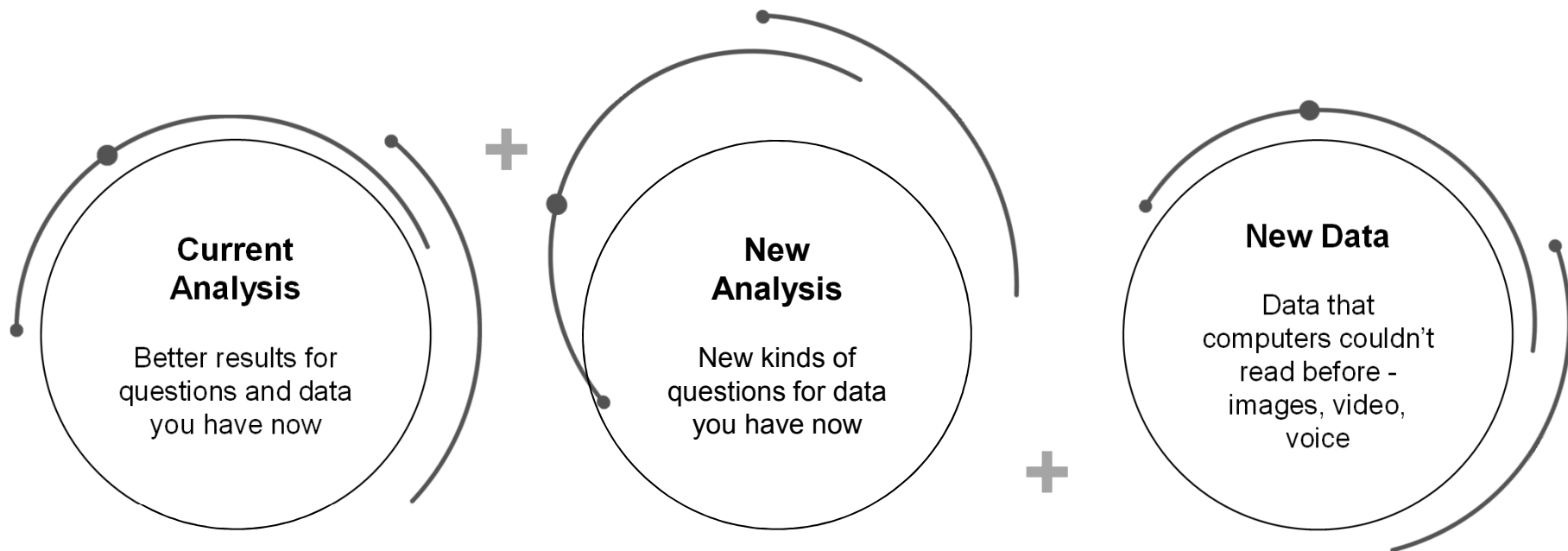
AI systems that augment human decision making and continuously learn from their interactions with humans and the environment.

Autonomous Intelligence

AI systems that can adapt to different situations and can act autonomously without human assistance



Why now (3/6)?
AI goes beyond improving current analysis

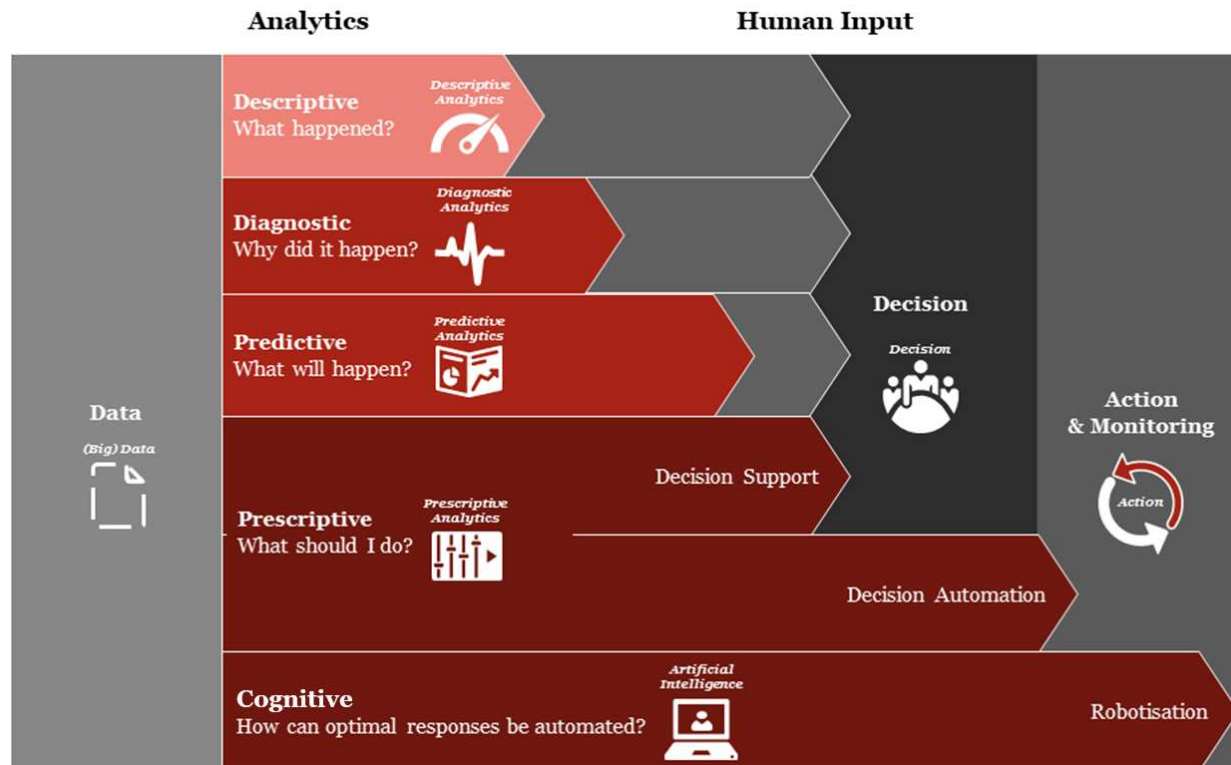


source: <https://www.youtube.com/watch?v=cVYDkPidXrU>



Why now (4/6)?

Maturity of data analytics increases to prescriptive/cognitive



Source: Gartner (2017)



Why now (5/6)?

Four key drivers have pushed advanced analytics from theory to reality



Accelerating Technology

Exponential increase in compute density



Big Data

Digitization of 'everything' creates lots of data



Open Source Software

Time to market drops as software is shared



Cloud Computing

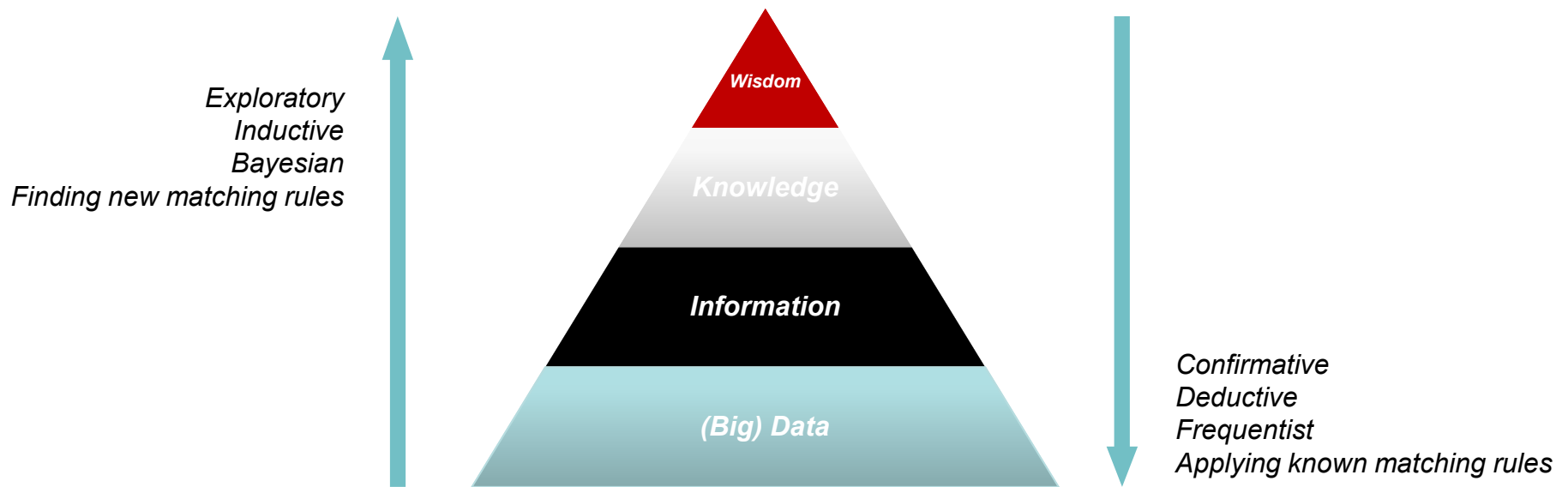
Cloud computing reduces marginal costs

Source: PwC (2018)



Why now (6/6)?

AI tends to focus less on deduction and more on induction



Source: Stephen Tuthill (1990)



Limpert Instituut

Met Bayesiaanse Statistiek het Kaf van het Koren Scheiden

Eric-Jan Wagenmakers
Psychologische Methodenleer
UvA



Limperg Instituut



Koen
Derks



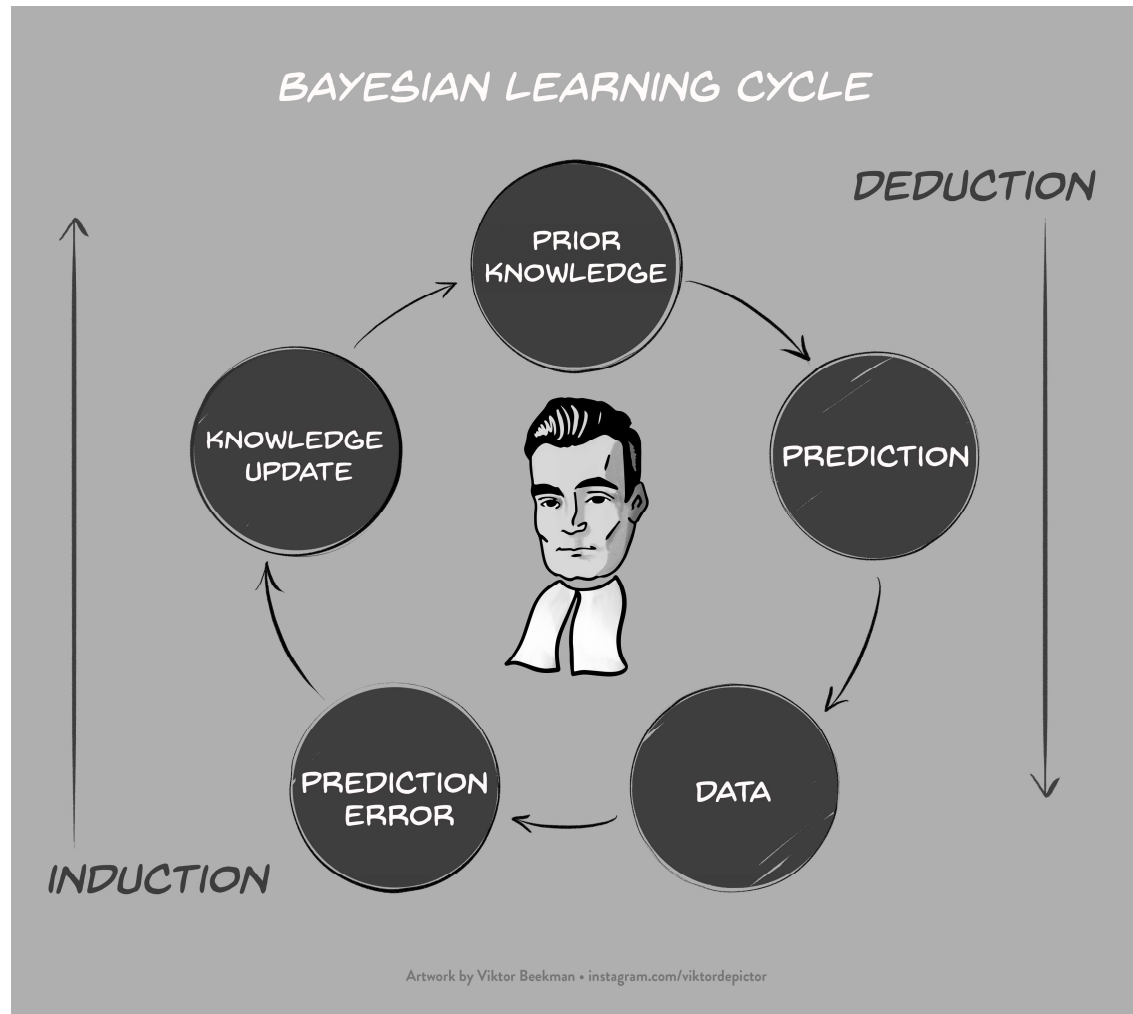
Ruud
Wetzels



Jacques
de Swart



Bayesiaanse Basis





Bayesiaanse Basis

- Aanpassen van onzekerheidsoordelen naar voorspellend succes: waardes die goed voorspellen worden waarschijnlijker, waardes die slecht voorspellen worden onwaarschijnlijker.



Bayesiaanse Basis

- Stel een accountant maakt claim C: “minder dan 5% van de totale begroting bevat een fout”.
- Wanneer heeft de accountant “bewijs” voor claim C?



Bayesiaanse Basis

- Wanneer heeft de accountant “bewijs” voor claim C?
- Als $P(C \mid \text{data}) > P(C)$
- Dat wil zeggen: de data hebben de claim waarschijnlijker gemaakt dan zij voorheen was.
- Als $P(C \mid \text{data}) < P(C)$ zal de accountant zich hopelijk hoogst ongemakkelijk voelen met C.

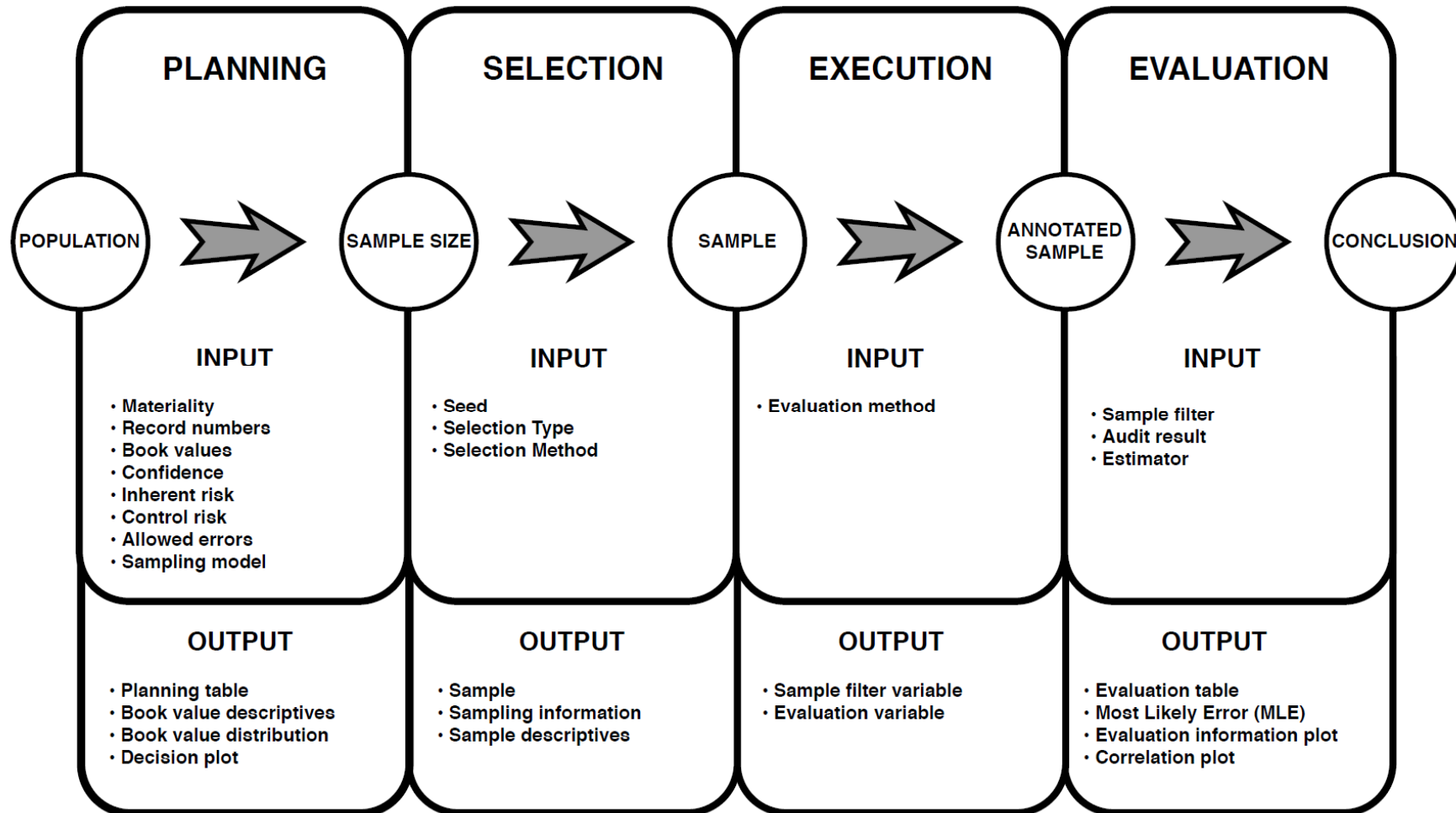


Bayesiaanse Basis

- Als $P(C \mid \text{data}) < P(C)$ zal de accountant zich hopelijk hoogst ongemakkelijk voelen met C.
- Echter: de klassieke/frequentistische benadering houdt zich niet bezig met bewijs in deze zin.
- In het vervolg betuig ik dat de Bayesiaanse benadering de moeite waard kan zijn, als toevoeging (of zelf vervanging!) op de huidige praktijk.



JASP for Audit





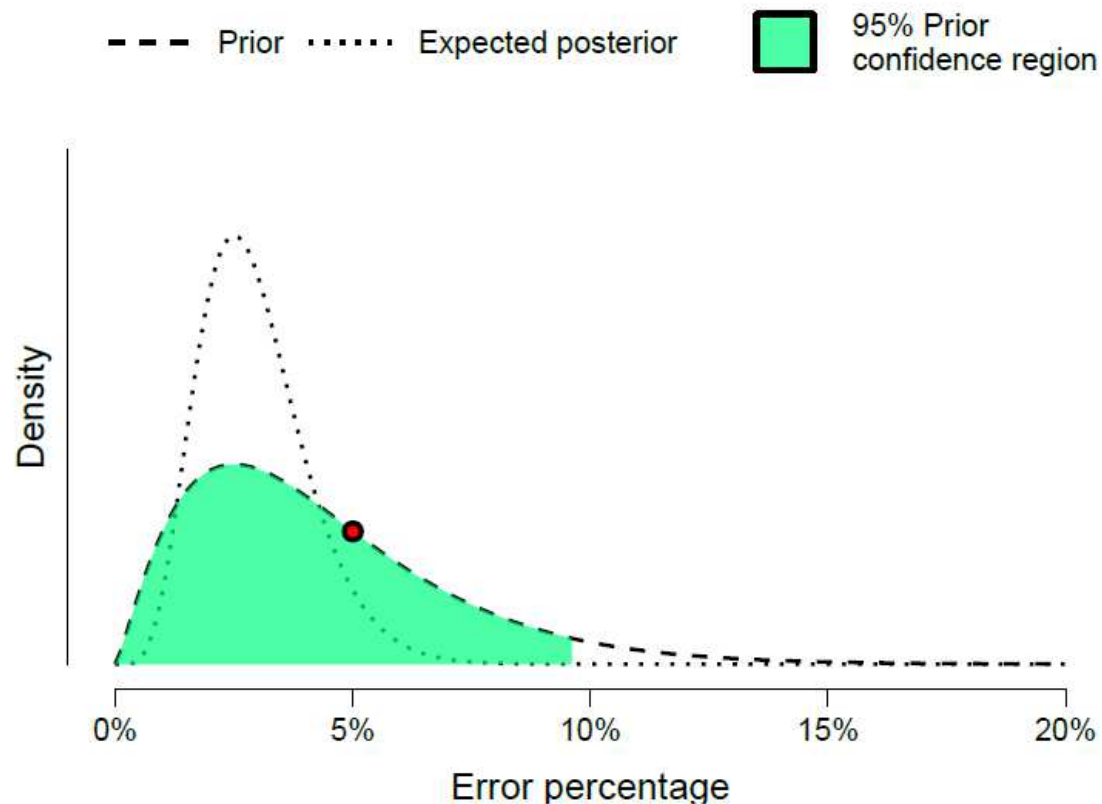
JASP for Audit

The screenshot shows the 'Audit Workflow' window in JASP. It is currently on the '1. Planning' step. Under 'Materiality', 'Absolute' is selected with a value of 0. Under 'Audit risk', 'Confidence' is set to 95%. The 'Variable selection' section shows a list of variables on the left: ID, bookValue, and auditValue. On the right, there are two empty input fields for 'Record numbers' and 'Book values'. Below these are expandable sections for 'Advanced options', 'Tables and figures', '2. Selection', '3. Execution', and '4. Evaluation'. At the bottom right of the workflow area are buttons for 'Download report' and 'To selection'.



Example of Planning

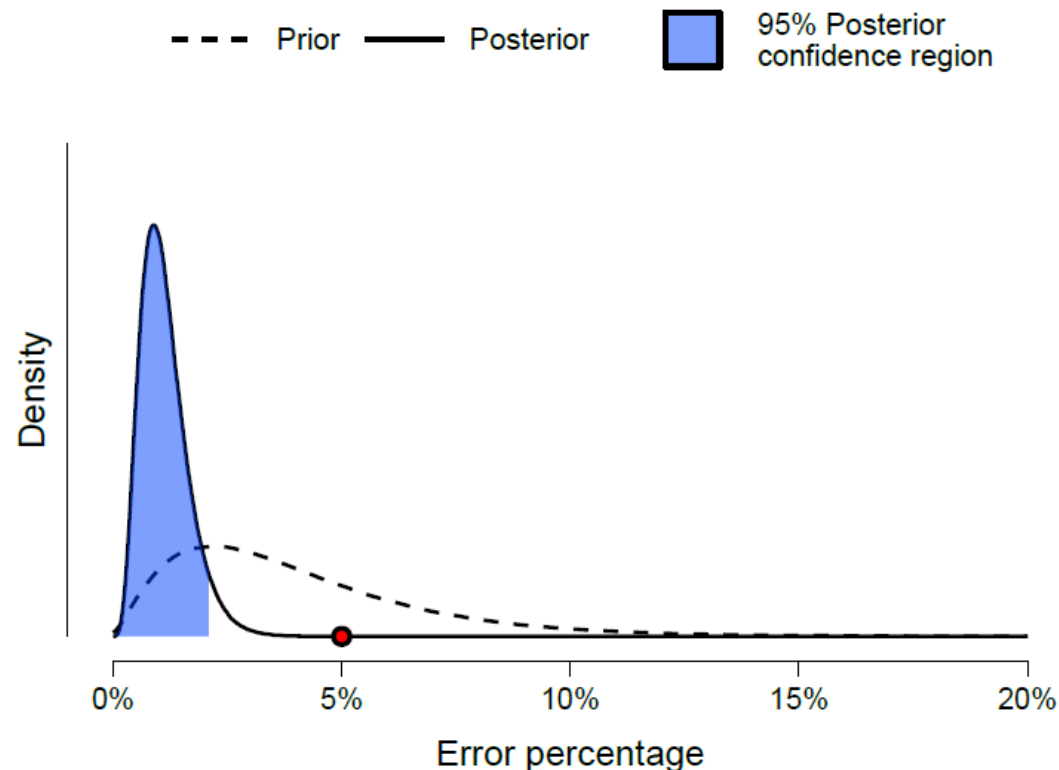
- Verwachting 2.5% fouten met medium control risk:





Example of Evaluation

- Na 3 fouten uit een steekproef van $n=169$:





Vragen

- Hoeveel maakt zo'n prior nu uit?
- Kunnen we ophouden met plannen en na iedere observatie onze kennis aanpassen?
- Of moeten we juist *meer* plannen, dwz *tijdens* het controleproces in plaats van alleen vooraf?



Limpert Instituut

Zelflerende algoritmen bij het verbeteren van datakwaliteit

W.J. Willemse (DNB)



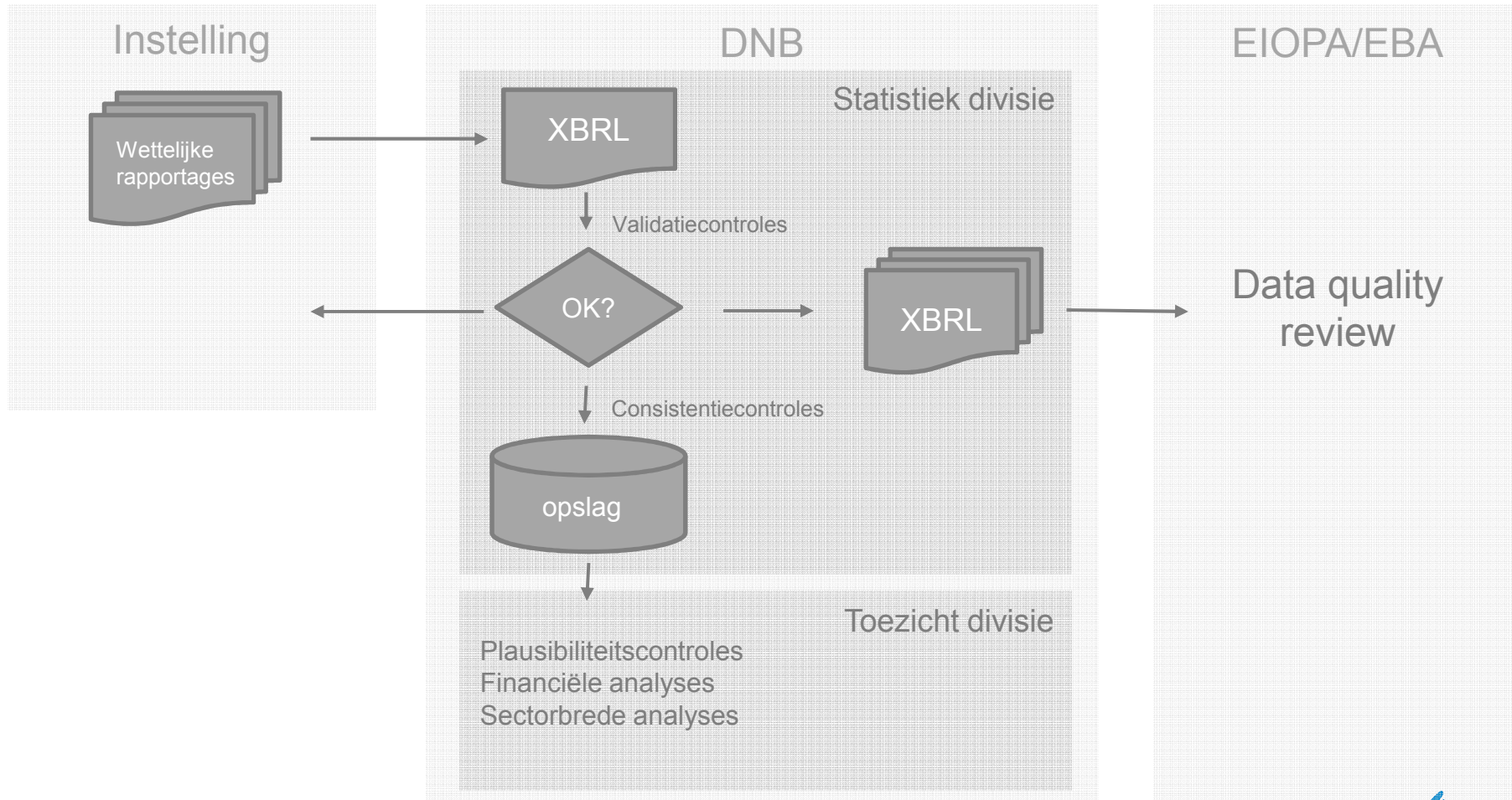
1. Toezichtrapportages bij DNB

- zeer veel datapunten per entiteit
- veel domeinkennis nodig bij interpretatie
- (te) veel werk om alle controles uit te schrijven

2. Zelflerende algoritmen bij plausibiliteitscontroles

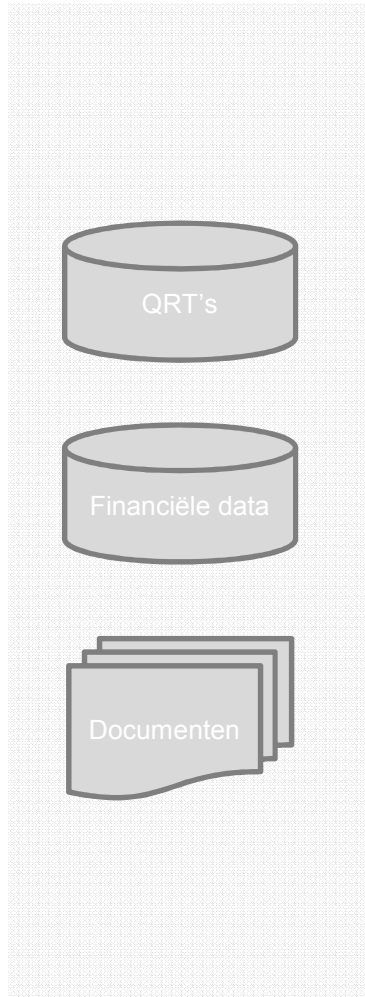
- in staat om verbanden en patronen in data herkennen zonder kennis vooraf
- genereren signalen die (kunnen) wijzen op datafouten, uitbijters en contra-intuïtieve veranderingen
- toepassing is beslissingsondersteunend

3. Uitdagingen bij toepassing open source software binnen organisaties





Limperg Instituut Data workflow plausibiliteitscontroles



4

Data clustering

Referentiegroepen
k-means, t-sne

PCA

Data analyse

Outlier detection

Patroonherkenning

Expert regels

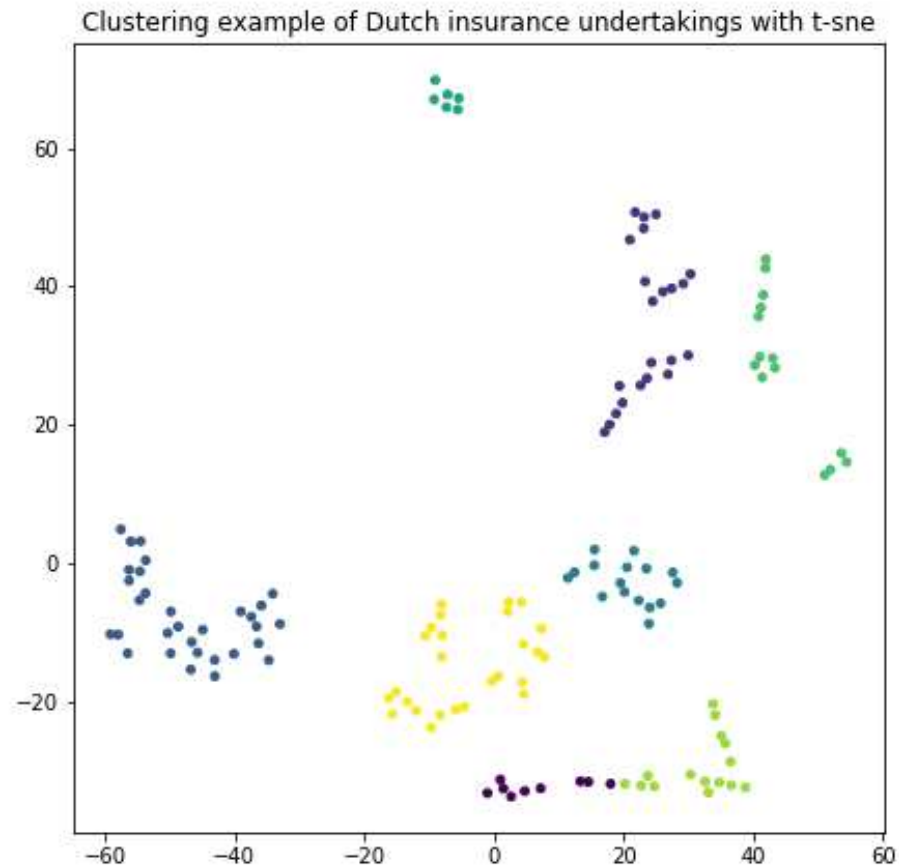
Resultaten

Signalen

Trendanalyse



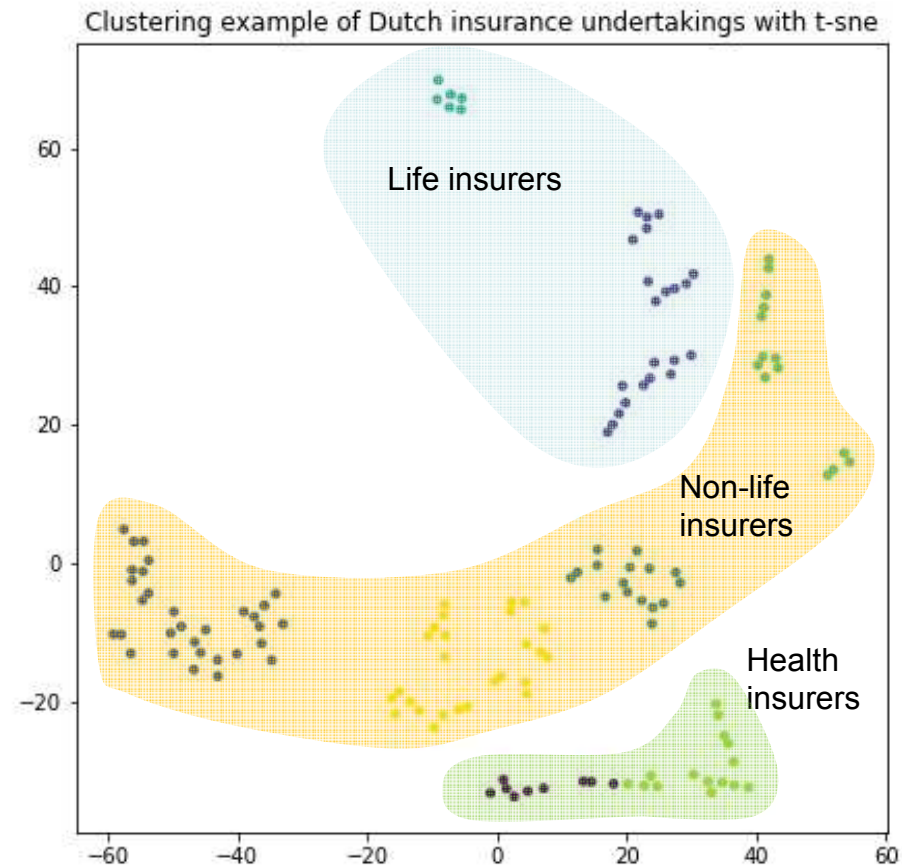
- Automatische classificatie met dimensiereductie en clustering
- Het t-sne algoritme vat n-dimensionale data samen in 2 dimensies met behoud van overeenkomsten en verbanden
- De afbeelding geeft het resultaat op basis van de balansen van Nederlandse verzekeraars (SII: 80 dimensies)





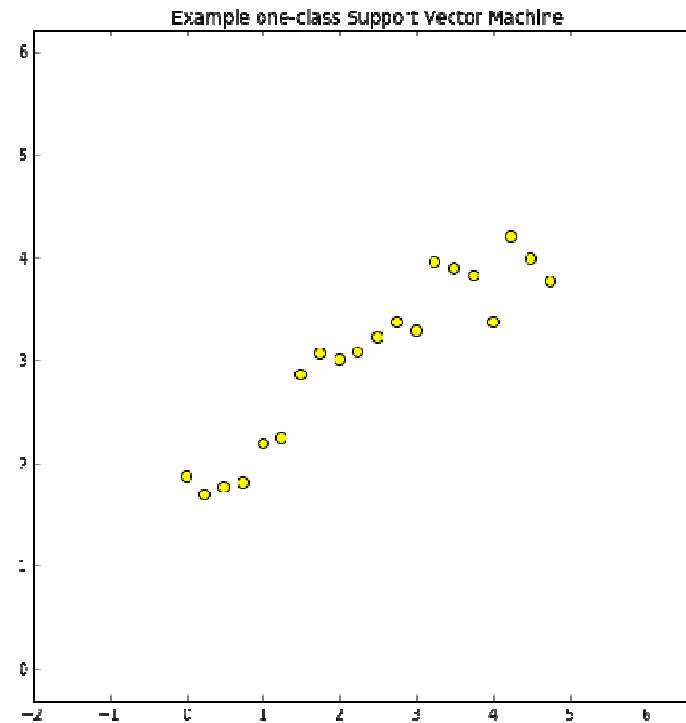
Limperg Instituut Dimensiereductie en clusteren

- Automatische classificatie met dimensiereductie en clustering
- Het t-sne algoritme vat n-dimensionale data samen in 2 dimensies met behoud van overeenkomsten en verbanden
- De afbeelding geeft het resultaat op basis van de balansen van Nederlandse verzekeraars (SII: 80 dimensies)



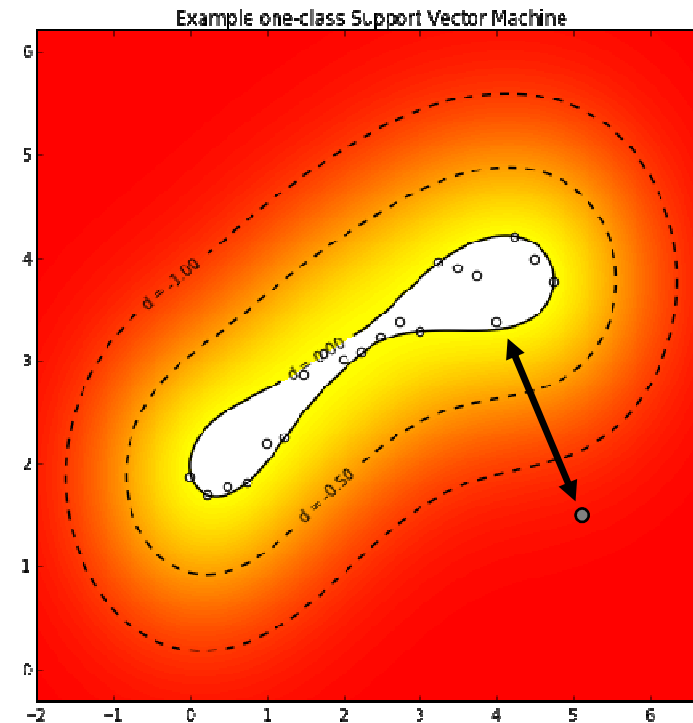


	X	Y
Data point 1	0.00	1.86
Data point 2	0.25	1.68
Data point 3	0.50	1.77
...
Data point 20	4.75	3.76



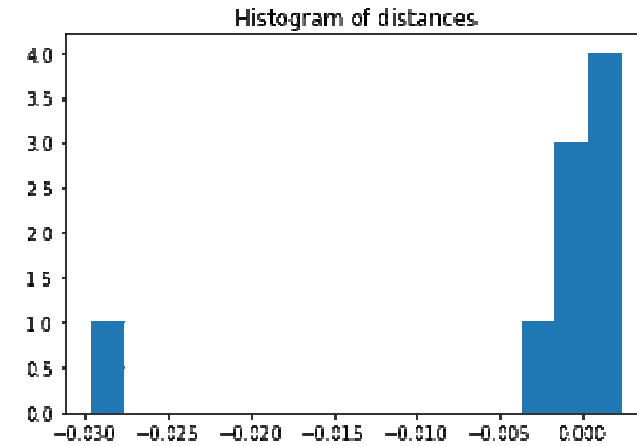


- Een Support Vector Machine bepaalt een *decision boundary* op basis van een gegeven *training set*, bijvoorbeeld historische rapportages
- Uitbijters in nieuwe rapportages worden herkend aan de hand van de afstand tot de *decision boundary*
- De afbeelding laat een 2-dim voorbeeld zien; het algoritme werkt voor n -dim data





Periode	CET-1 vermogen	Total capital ratio	T1 capital ratio	CET-1 capital ratio
2014H2	40543	14,6%	13,5%	13,5%
2015H1	39874	17,0%	14,3%	12,8%
2015H2	41554	16,9%	14,5%	12,9%
2016H1	41986	18,0%	15,1%	13,2%
2016H2	44466	19,3%	16,3%	14,2%
2017H1	44888	19,6%	16,3%	14,5%
2017H2	45581	18,5%	16,2%	14,7%
2018H1	44833	18,4%	15,7%	14,1%
Nieuwe rapportage (hypothetisch)				
2018H2	35000	18,2%	9,7%	13,1%



(Op basis van: individuele gegevens banken (jaar), statistiek.dnb.nl)



Limperg Instituut Patroonherkenning - associatieregels (1)

- Een associatie regel beschrijft de associatie tussen twee delen in een dataverzameling of tussen dataverzamelingen

	Shopping list
1	bread, milk, eggs
2	bread, milk, eggs, bacon, cheese, muesli
3	milk, muesli
4	bread, eggs

- De associatie {melk} → {muesli} heeft
 - Een ondersteuning / support van 2 / 4 (twee van de vier lijstjes hebben zowel melk als muesli)
 - Een betrouwbaarheid / confidence of 2 / 3 (twee van de drie lijstjes met melk heeft ook muesli)



Limperg Instituut Patroonherkenning - associatieregels (2)

- Binnen toezichtrapportages zoeken we naar patronen zoals samenhangende datapunten in de rapportages en verbanden met eerdere rapportages
- Het associatie-algoritme ontdekt regels; vervolgens kijken we naar data die de regels bevestigen of juist met de regels in strijd zijn

- Voorbeelden:

Support	Confidence	Rule
134	96%	<code>risk_margin ≤ technical_provisions</code>
132	95%	<code>assets held for unit-linked contracts ≤ technical provisions - unit-linked</code>
122	88%	<code>net scr - market risk = gross scr - market risk</code>
93	89%	<code>delta(reinsurance recoverable from non-life) ≥ 0 and delta(technical provisions - non-life) ≥ 0</code>

(Op basis van: individuele gegevens verzekeraars (jaar), statistiek.dnb.nl)



Limperg Instituut **Uitdagingen bij open source software**

- Gebruik van open source software vereist aandacht
 - noodzaak tot gestructureerd beheer van omgevingen en codeversies
 - afspraken nodig om reproduceerbaarheid te waarborgen
 - ondersteuning (zelf) organiseren
- Implementatie in bestaande processen en infrastructuur niet altijd eenvoudig
 - uitlegbaarheid en transparantie
 - feedbackloop met gebruikers
 - toegang tot ongestructureerde data



Limpert Instituut

Kunstmatige Intelligentie ontrafeld

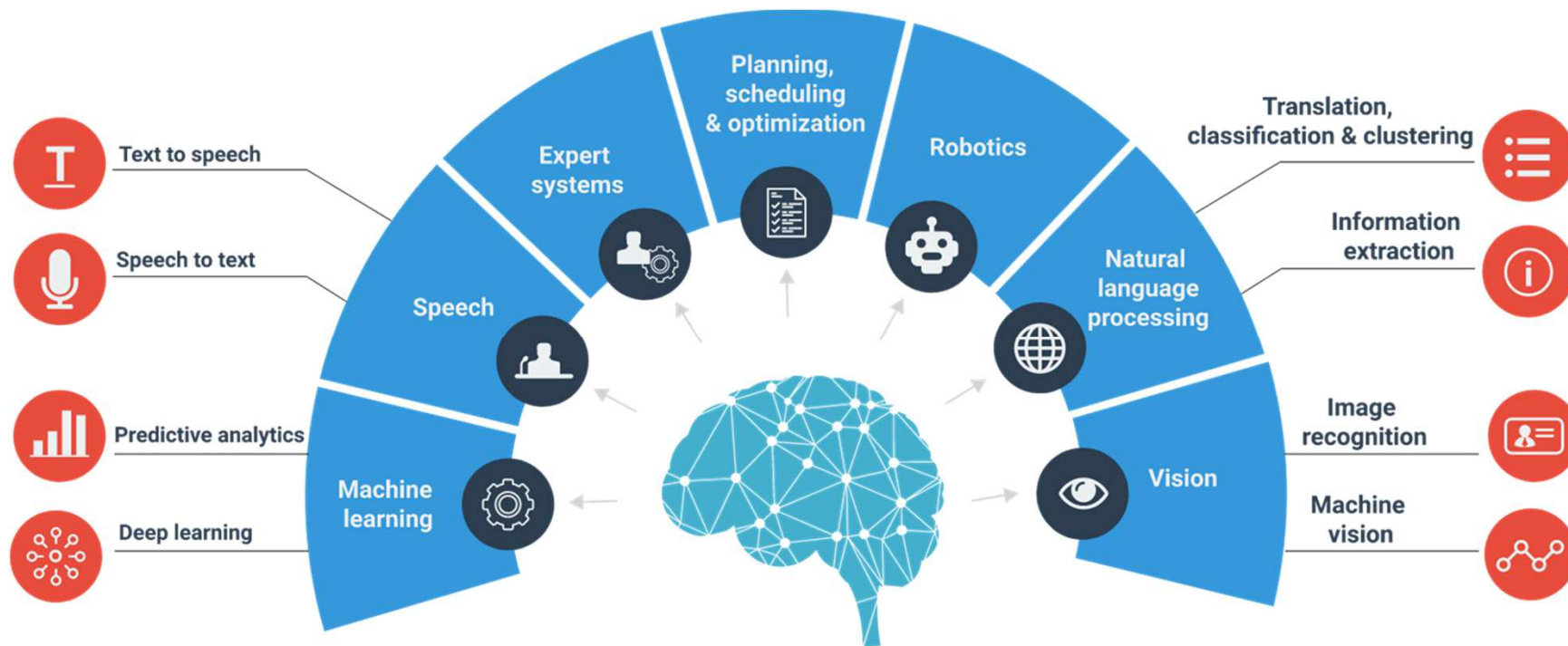
Dr. Tjerk de Greef
Director of Search Technology



Wolters Kluwer



Kernbegrippen AI



Machine learning is the science of getting computers to act **without being explicitly programmed**

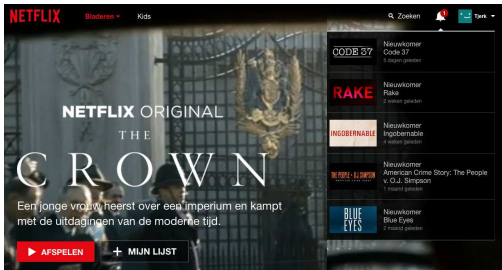
(Andrew Ng, Stanford University)



AI Toepassingen

Persoonlijke Aanbevelingen

Spotify, Amazon, Netflix



Beeldherkenning

Facebook, Apple, Google, securities

Who's in These Photos?

The photos you uploaded were grouped automatically so you can quickly label and notify friends in these pictures. (Friends can always untag themselves.)



Robotica



Swift key



Zelfrijdende Auto'scar



Spam filters





Spam?

hi intimate,

WARNING: Wanna miss the opportunity, then move on

VIAGRA, OTHER MEDICATIONS

Are you tired of giving heavy sum of money for prescription medications?

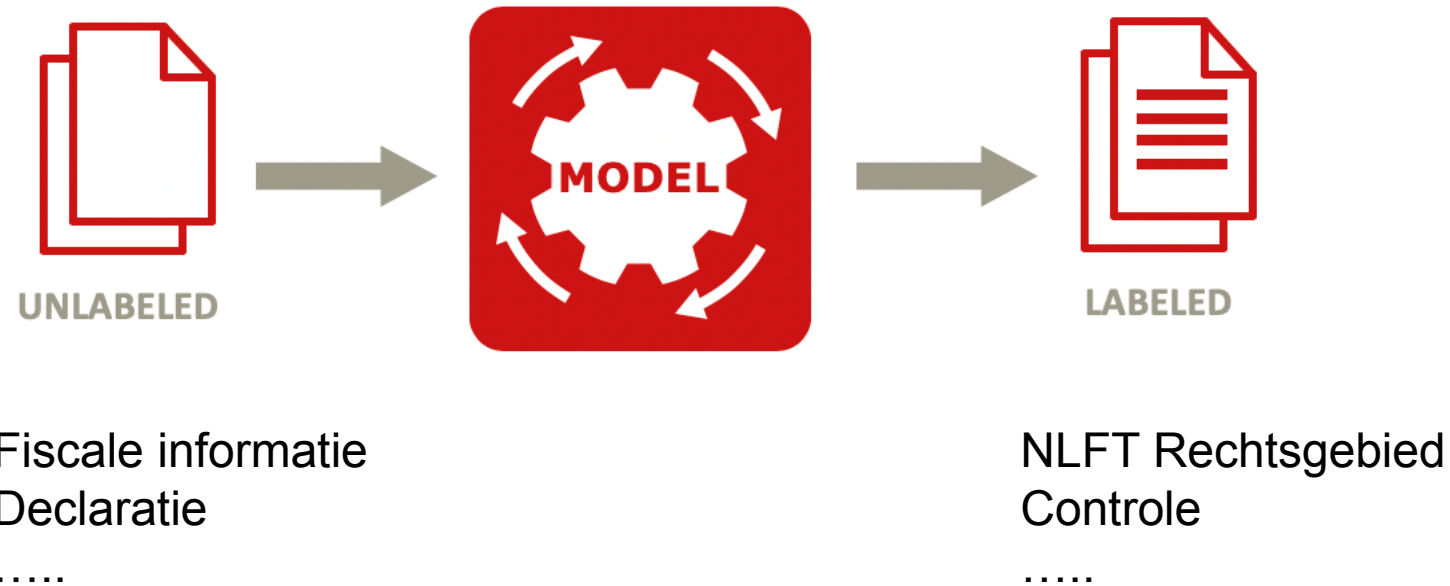
You can save up to 80% on all popular meeds.

Don't miss the opportunity!!! FREE



ML Aanpak (1)

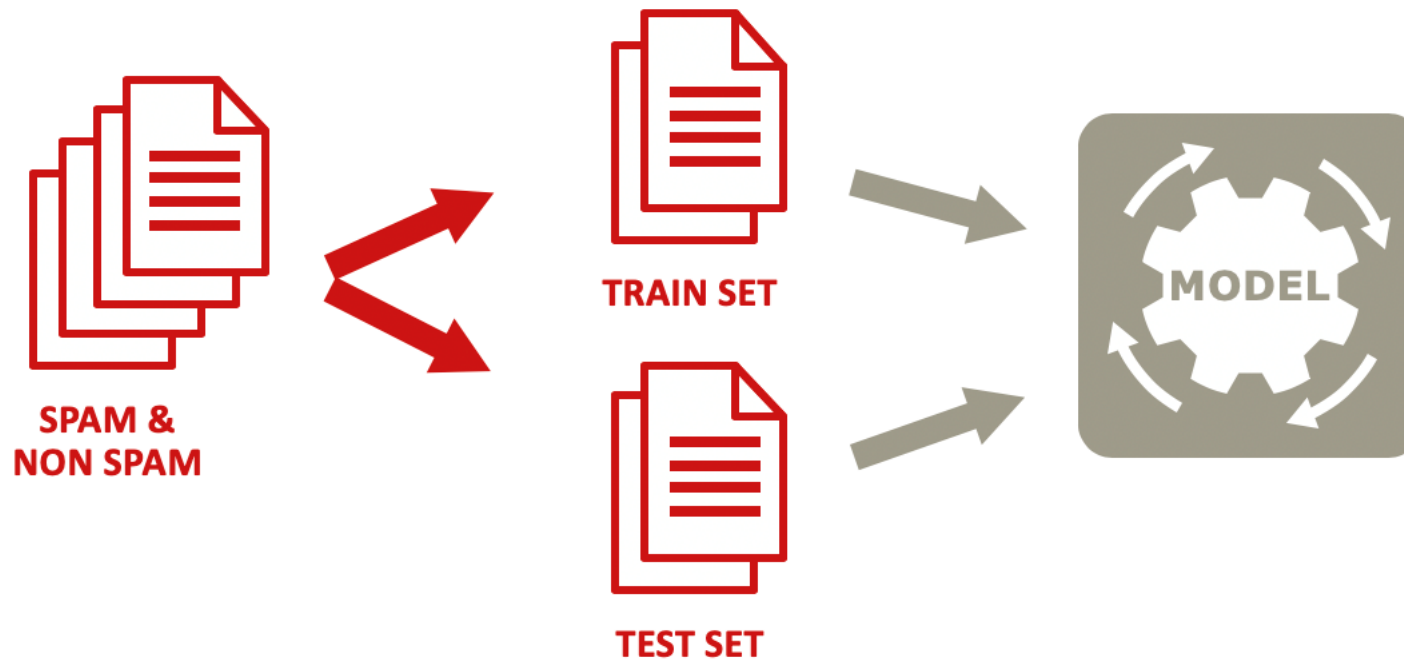
DOEL: VOORSPELLEND MODEL





ML Aanpak (2)

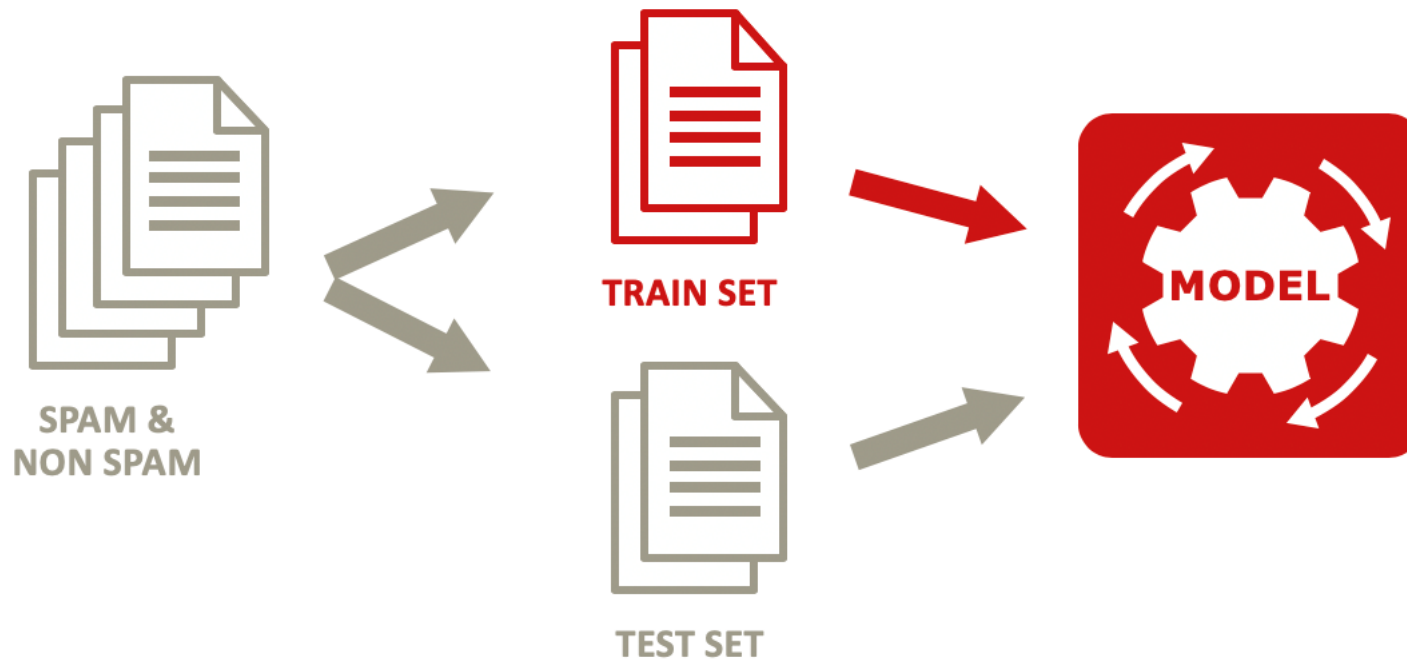
SPLITSEN DATA





ML Aanpak (3)

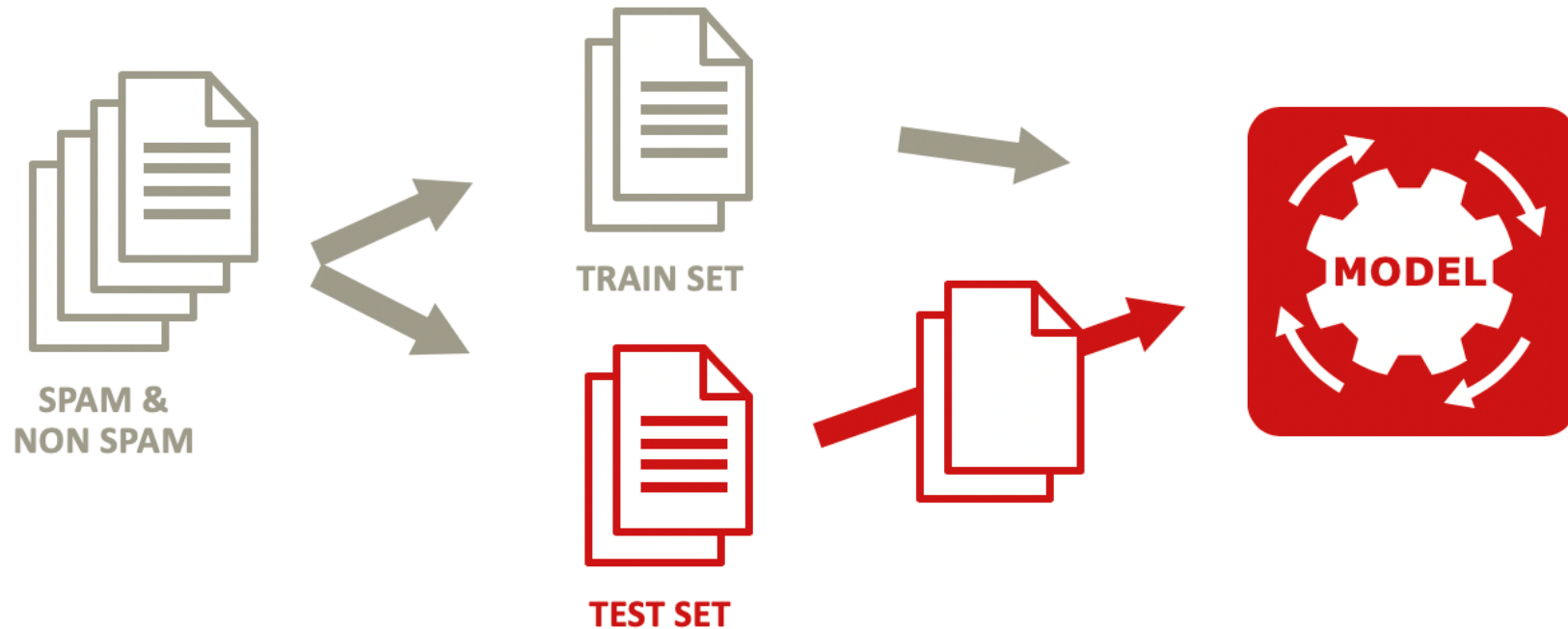
TRAIN MODEL





ML Aanpak (4)

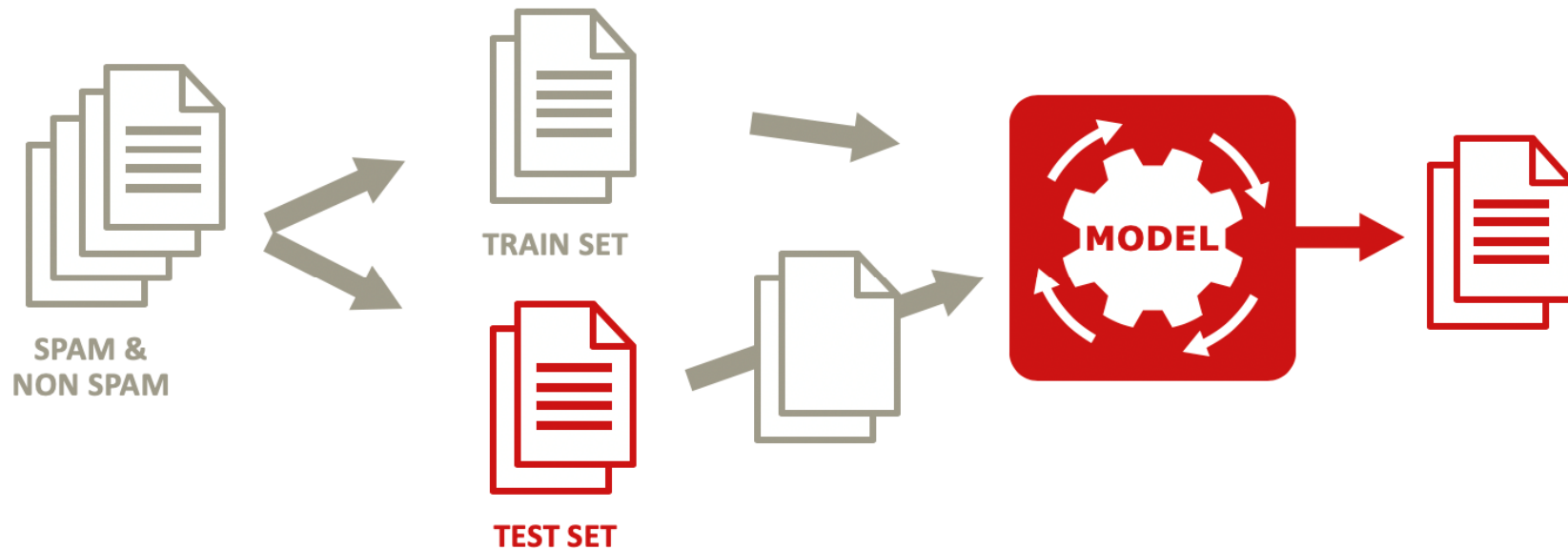
ONTLABEL TEST SET





ML Aanpak (5)

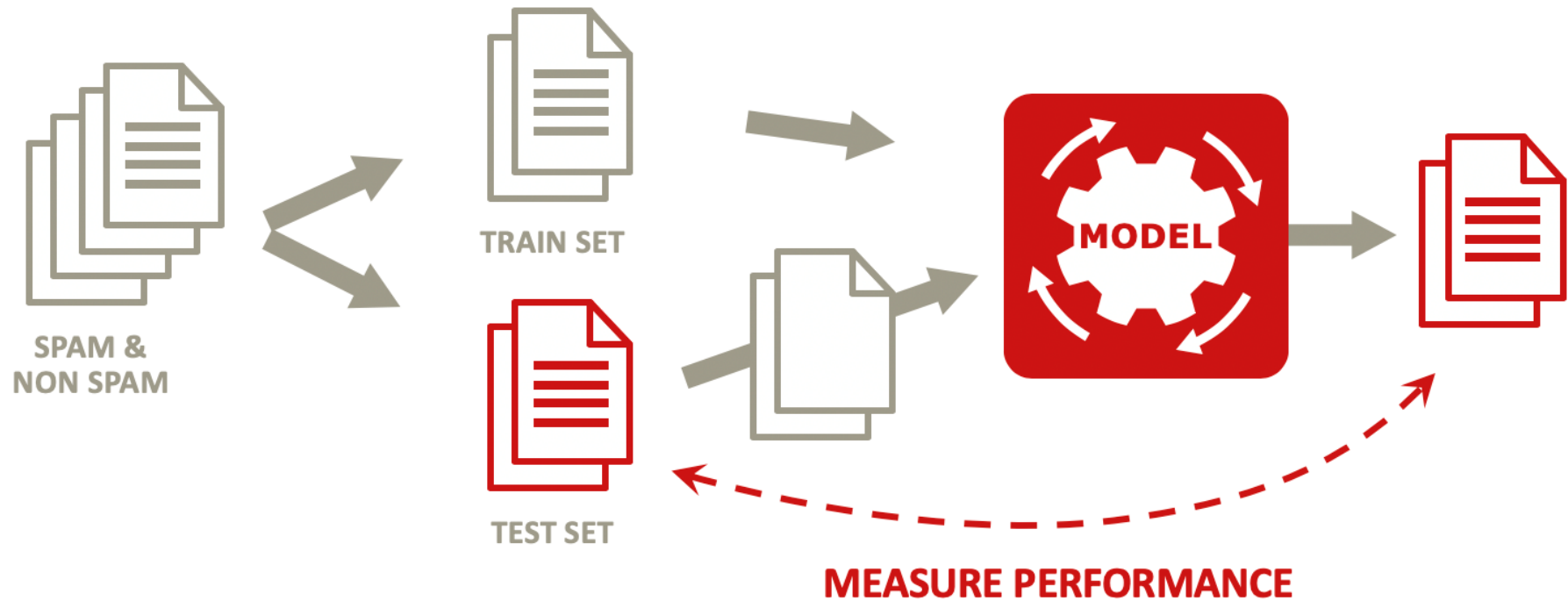
LAAT MODEL VOORSPELLEN





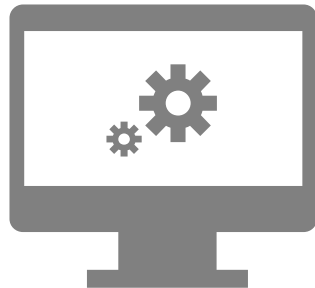
ML Aanpak (6)

EN VERGELIJK



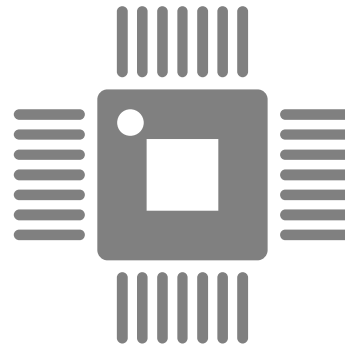


Waarom nu?



Software

- Efficiënte Programmeertalen
- Active Open Source Community
- Hot topic in zowel academisch als commerciële wereld



Hardware

- Cloud computing minimaliseert investeringen en biedt flexibiliteit
- Hergebruik van **GPU's** maakt matrixberekeningen bijzonder efficiënt (lineaire algebra)
- Deep Blue kost nu \$2,000 kosten i.p.v. \$15 miljoen

```
10010010011100100
10010001010010001
01010110101011010
10010010011100100
10010001010010001
01010110101011010
```

Data

- De hoeveelheid data groeit exponentieel
- Data krijgt waarde (Bureau van Dijk = \$3B)
- Steeds meer gestructureerde data maar ongestructureerde data kan steeds beter verwerkt worden



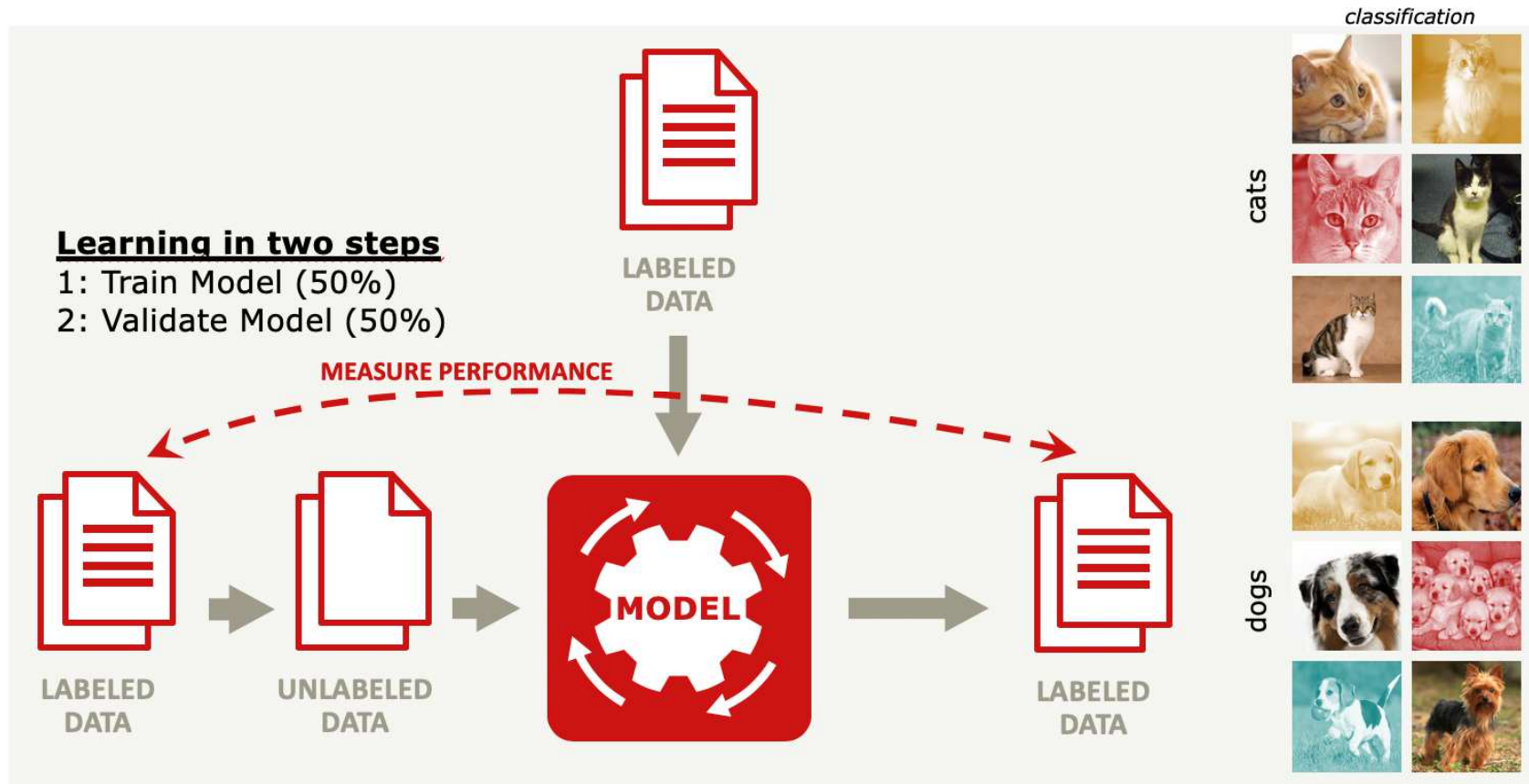
5 vormen van ML

1. Supervised Learning
2. Unsupervised Learning
3. Reinforcement Learning
4. Deep Learning
5. Natural Language Processing (NLP)



1. Supervised Learning

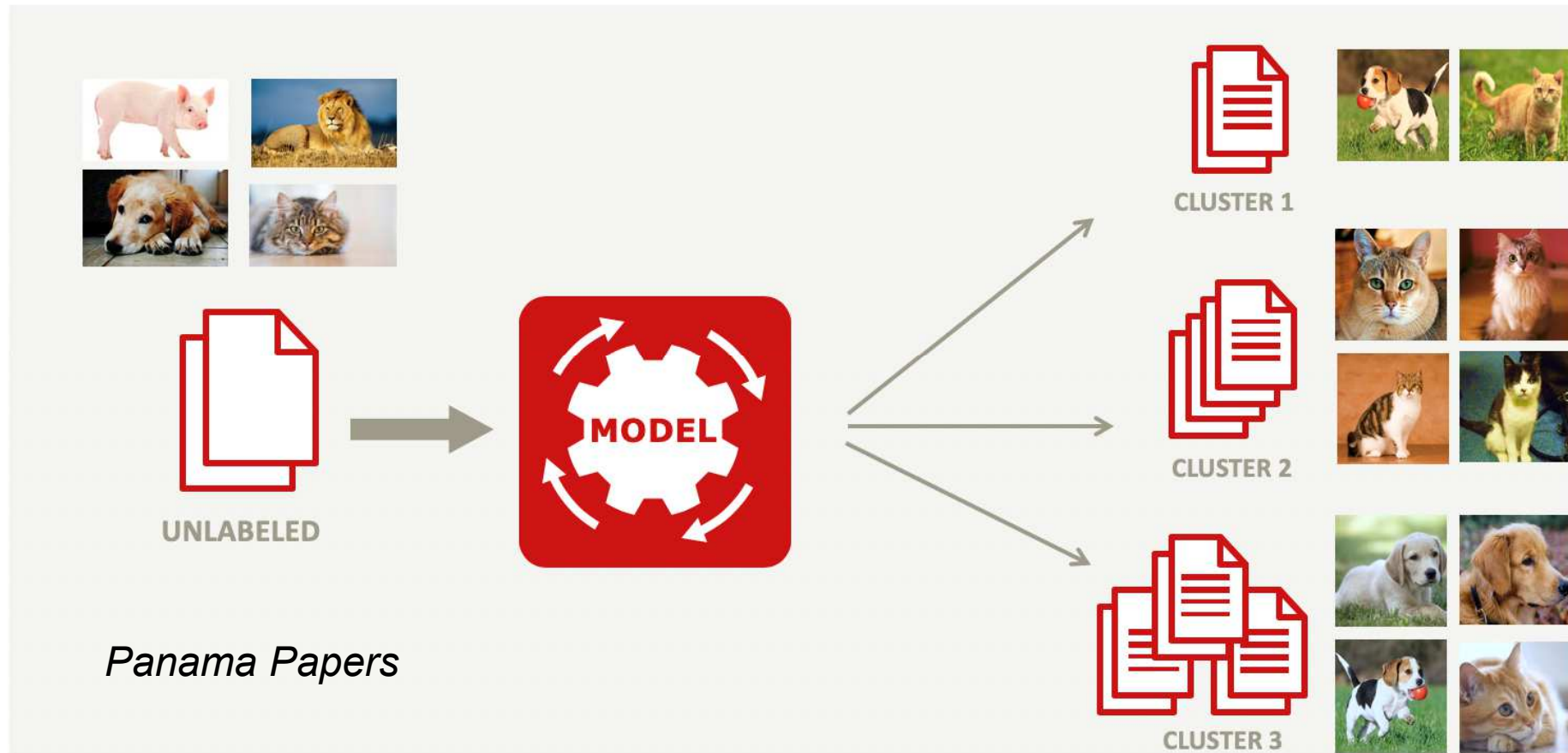
Structuur vinden m.b.t. **gelabelde** data





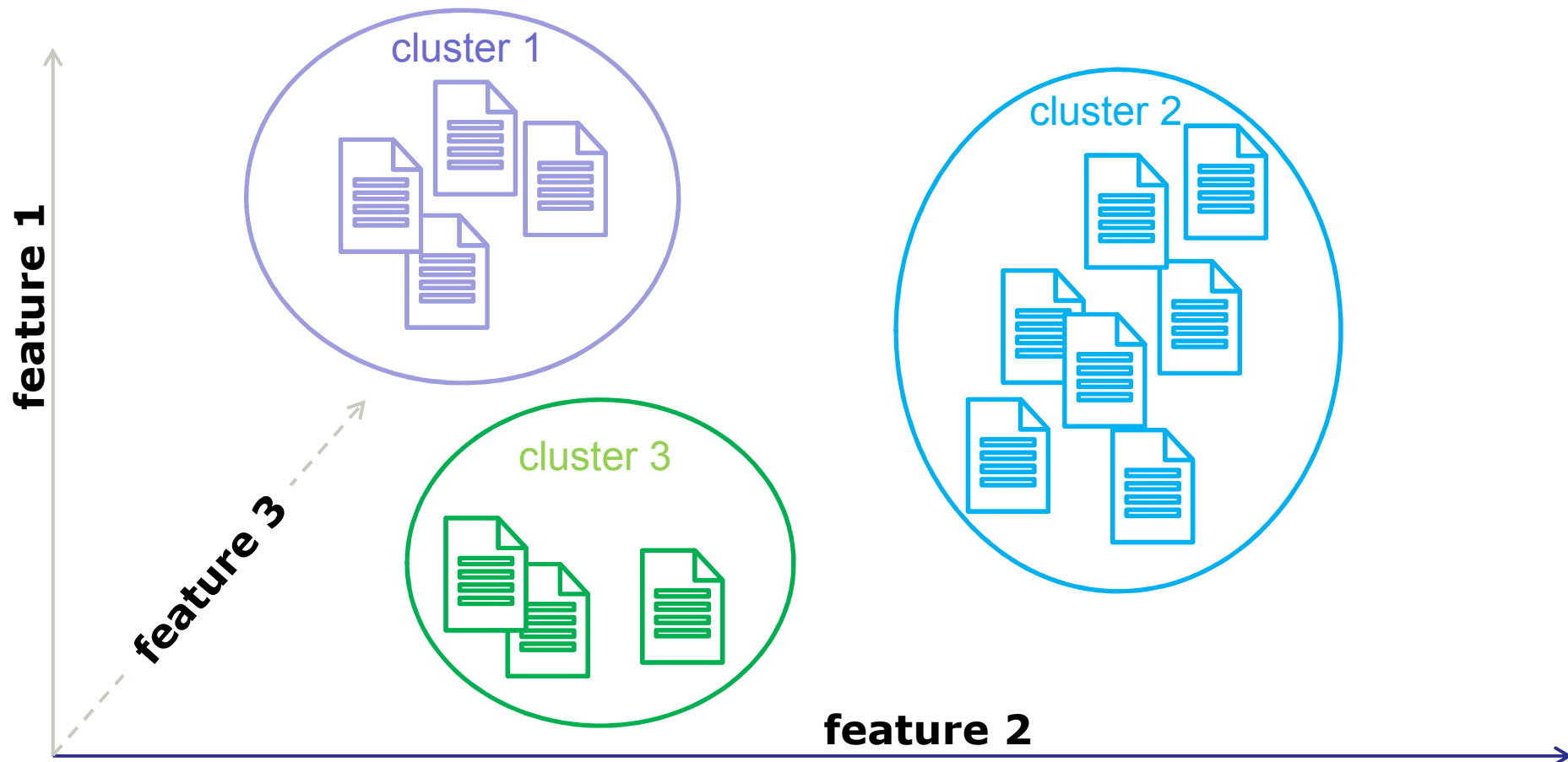
2. Unsupervised Learning

Structuur vinden m.b.t. **ongelabelde** data





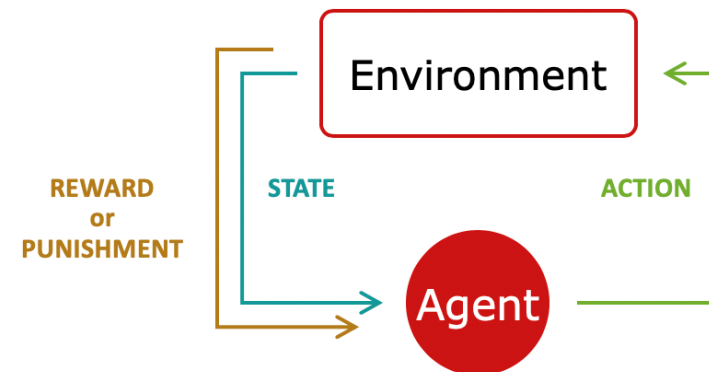
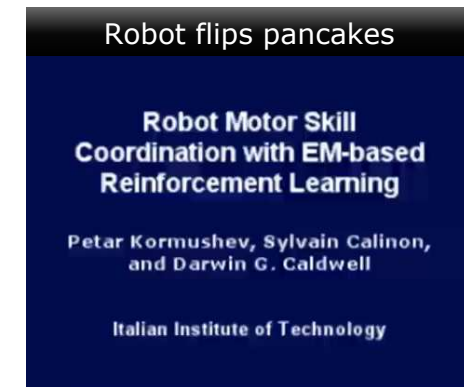
Leren = Clusters Vinden





3. Reinforcement Learning

- Geïnspireerd op psychologie
(operant conditioning)
- Twee soorten feedback
positief (beloning)
negatief (straf)



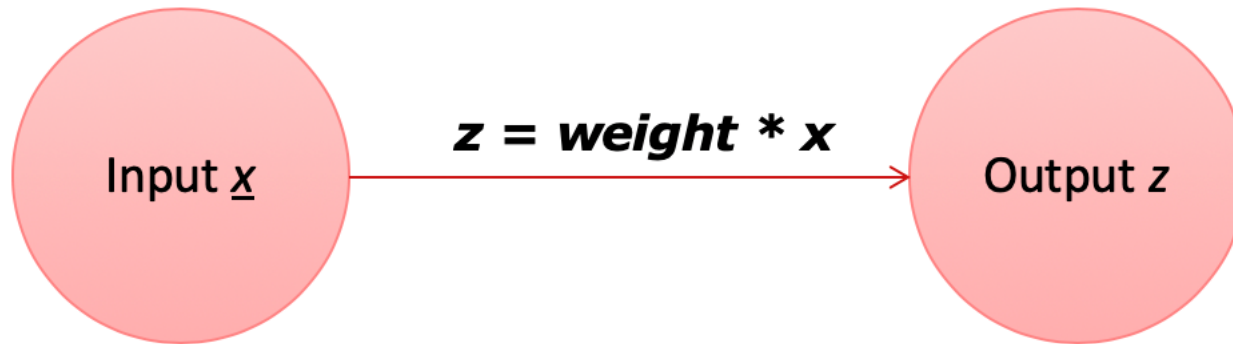
reward if clicked → 1
reward if not click → 0



4. Deep Learning

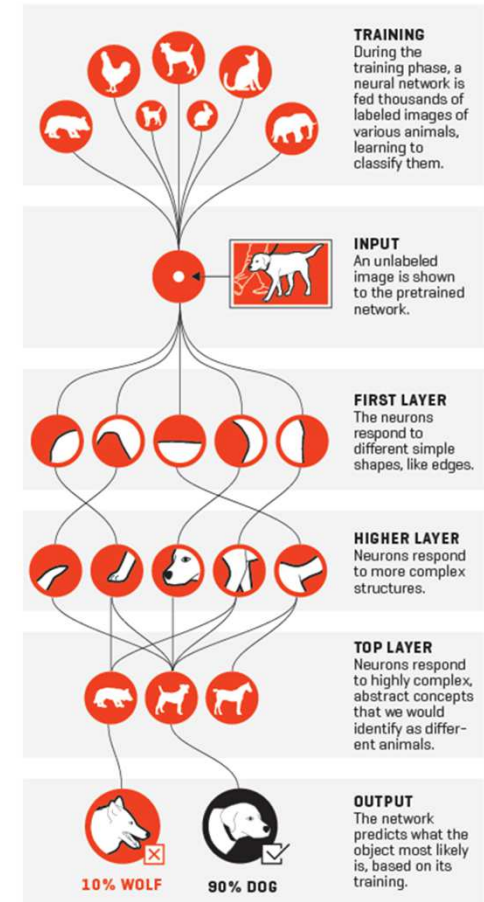
Structuur vinden m.b.t. **gelabelde** data door probleem in lagen op te splitsen

Learning Model is about adapting weights



in a large network of many layers deep

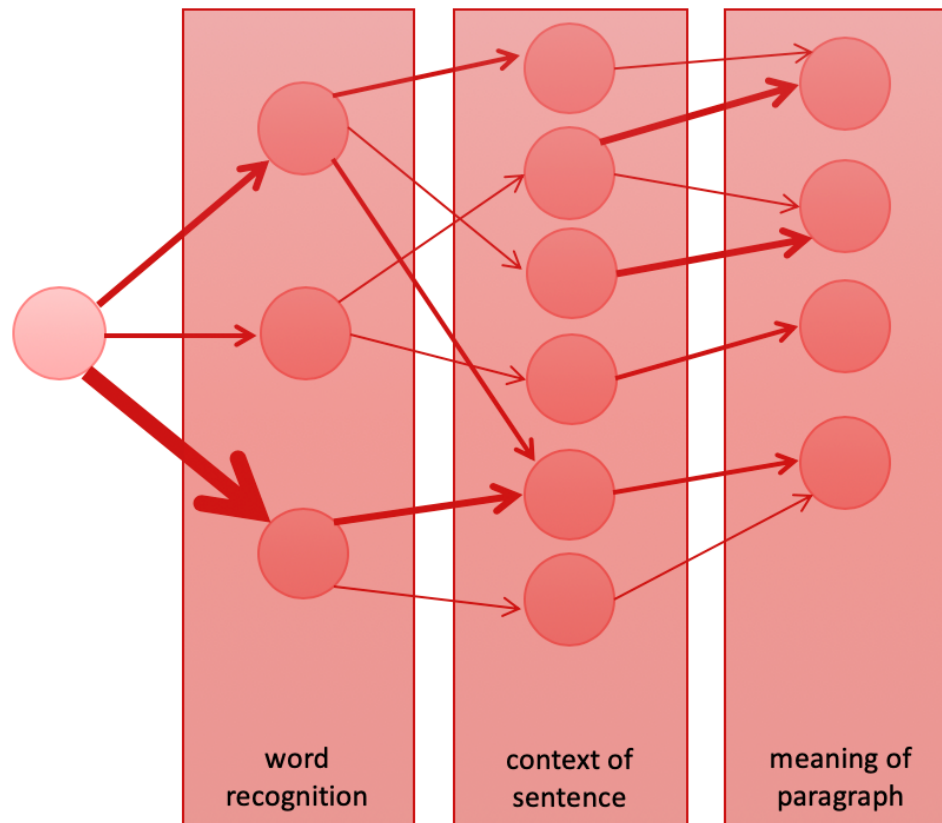
HOW NEURAL NETWORKS RECOGNIZE A DOG IN A PHOTO





4. Deep Learning

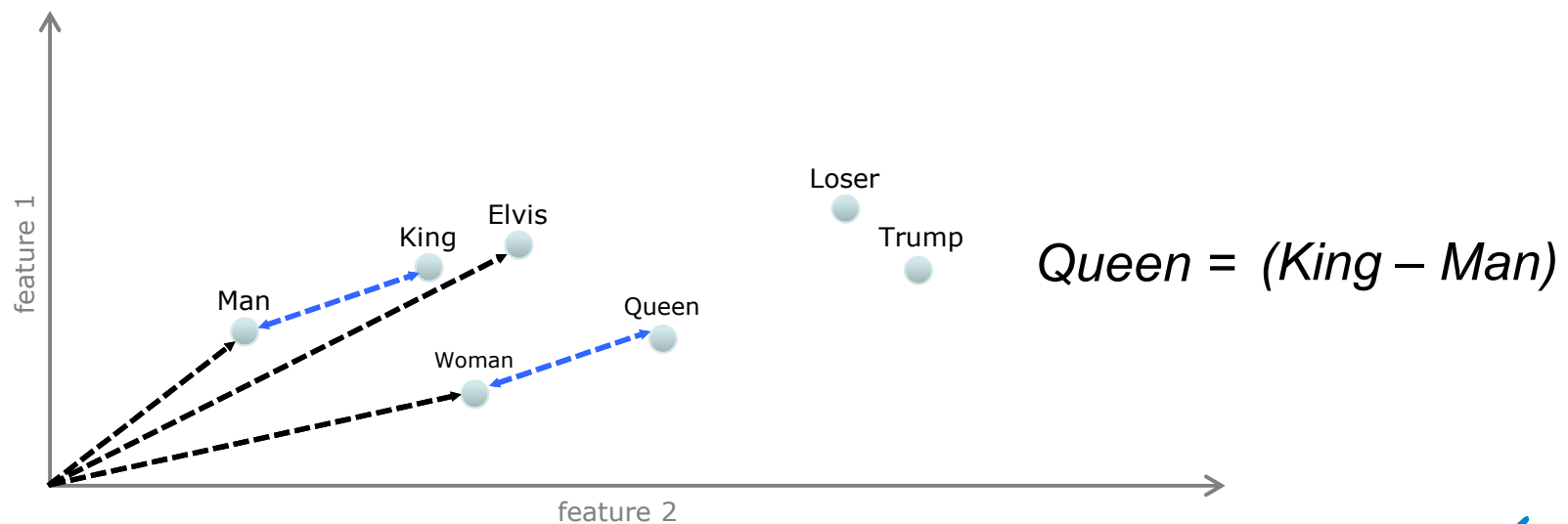
Structuur vinden m.b.t. **gelabelde** data
door probleem in lagen op te splitsen

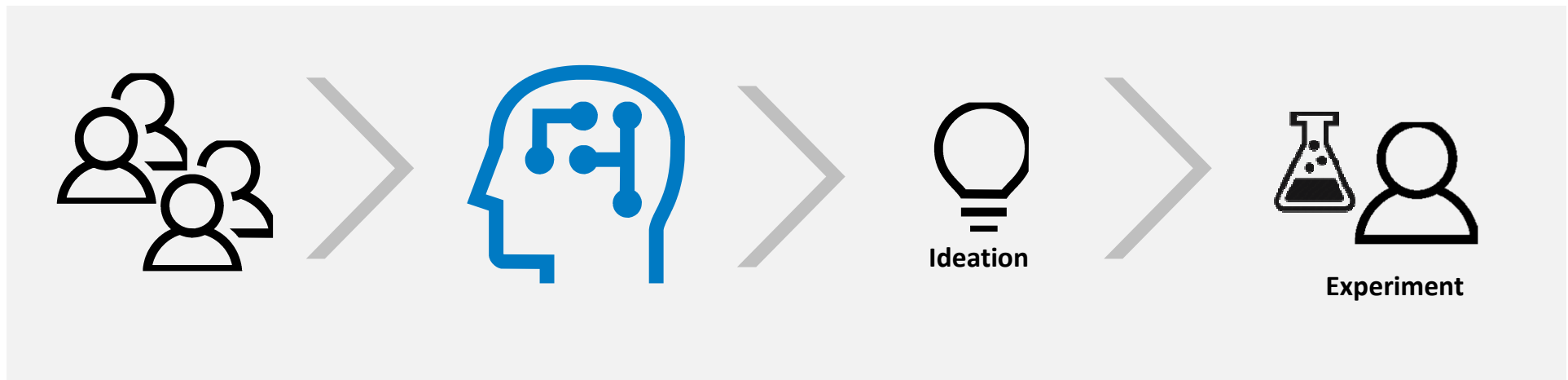




5. NLP

- Vroeger
 - Regels
 - Grammatica's
- Nu m.b.v.. Machine Learning (Word-2-Vec)





Begrijp AI

Waardenken

- AI Capabilities** 
- Cognitive Insights
 - Outlier Detection
 - Cognitive Engagement



Limperg Instituut

Vragen