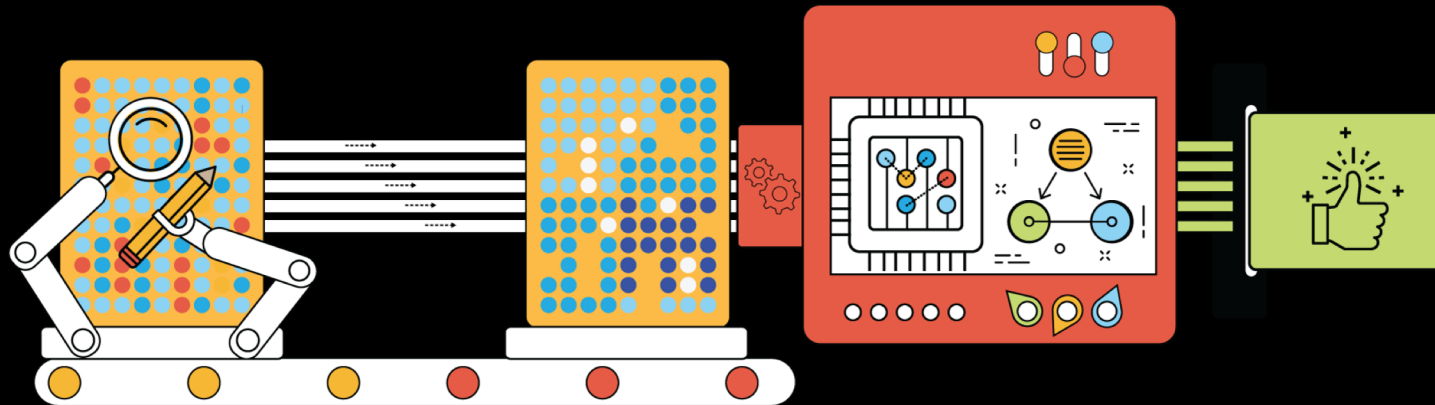


Datacoalitie Datagedreven onderzoek

Studenten nóg beter begeleiden met een datagedreven aanpak!



Agenda

1 Wie zijn wij

2 Verwachtingenmodel uitval niveau 2 student

3 Integraal Teamdashboard en onderzoek naar datakwaliteit

4 Verbeteren databeschikbaarheid werving, instroom en doorstroom

5 Meedoen?

11 mbo-instellingen onderzoeken samen met OCW en DUO hoe data kan worden ingezet om de studieloopbaan van studenten te verbeteren en wat de toegevoegde waarde is om dit met een datagedreven aanpak te doen.

11 mbo-instellingen:



Noorderpoort

roc van
twente



curio

ROC NIJMEGEN

gildeopleidingen

DUO:



Dienst Uitvoering Onderwijs
Ministerie van Onderwijs, Cultuur en
Wetenschap

OCW:



Ministerie van Onderwijs, Cultuur en
Wetenschap

Projectorganisatie:

hutspot

new
nodes

Wie zijn wij

Waarom zijn wij een datacoalitie?

- Samen leer je meer dan alleen
- Er zijn inhoudelijke (onderwijs) en data-experts nodig voor meer datagedreven werken
- Wij willen niet het wiel opnieuw uitvinden, maar leren van elkaar en daarop verder bouwen
- Data beschikbaarheid en analyse kennis en capaciteit van DUO is essentieel
- Resultaten, leerervaringen en technieken zijn beschikbaar voor de hele sector
Zie: <https://datagedrevenonderzoekmbo.nl/>

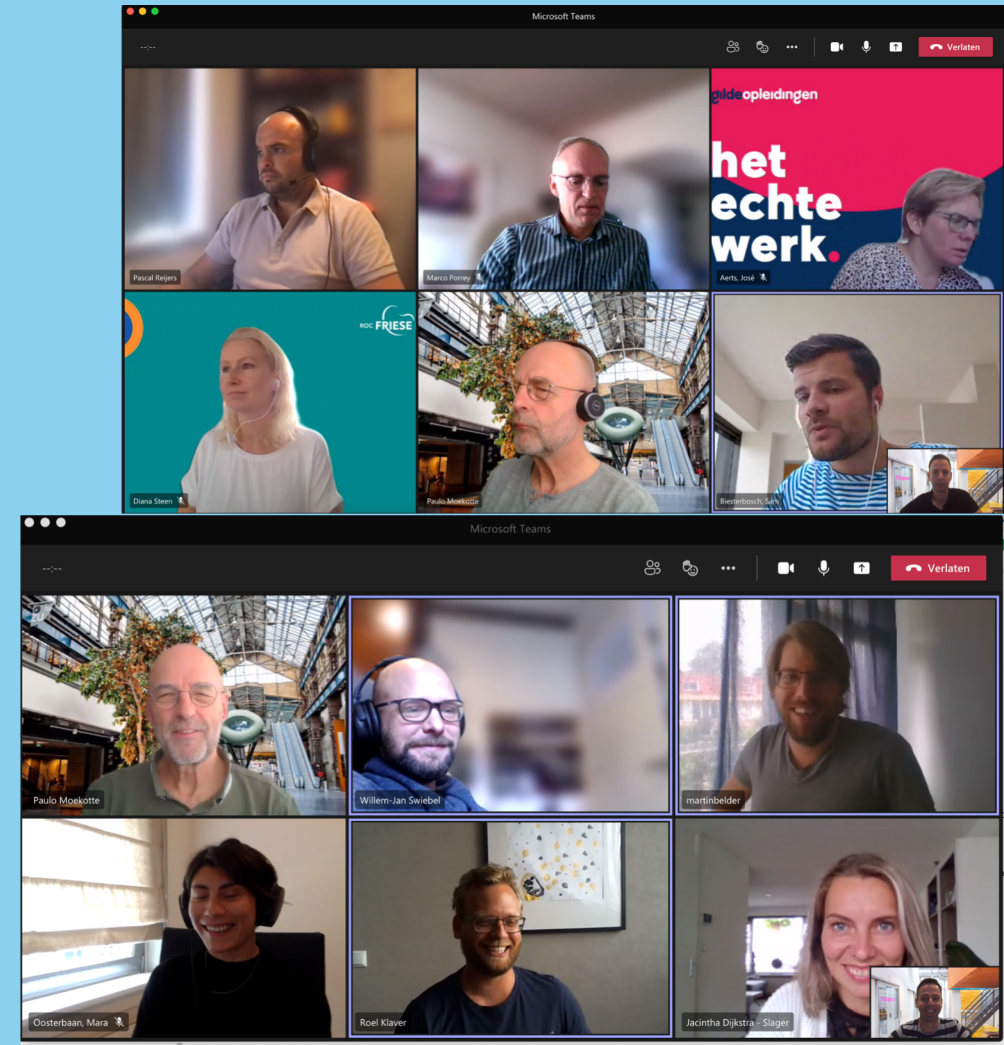
Wie zijn wij

Organisatie, kennis en ervaring

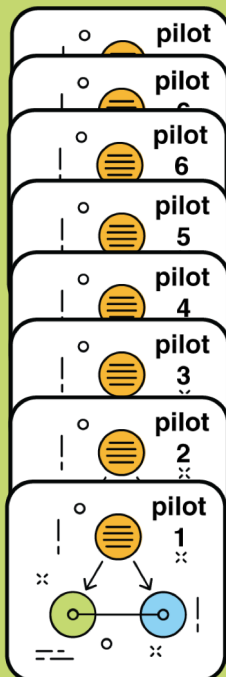
- Stuurgroep: bestuurlijke vertegenwoordiging per instellingen + dir. OCW en DUO
- Projectgroep: club van inhoudelijk experts
 - 1 of 2 vertegenwoordigers per instelling
 - Totaal hebben ca. 90 mensen actief bijgedragen aan het initiatief
 - Zowel vanuit onderwijsteams als vanuit informatiemanagement, ICT, control, kwaliteit, beleid en juridische zaken
- DUO is vanaf de start een belangrijke partner:
 - Toegang tot en brede beschikbaarheid van data (leerloopbaan student: PO-VO-MBO-HBO)
 - Kennis, expertise en analyse capaciteit
 - Middelen: de inzet van DUO is gefinancierd door OCW (bijdrage ontwikkeling datagedreven werken in het mbo)

Wie zijn wij

- 3 werkgroepen, 21 sessies per sprintfase,
- “244x je microfoon staat nog uit”
- Samenwerking van Noord tot Zuid
- Om datagedreven werken verder te brengen



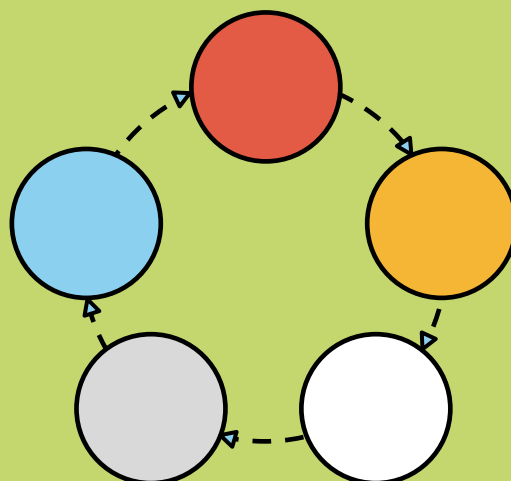
Bereikte resultaten



12 pilots

..waarbinnen onderzoek is uitgevoerd

levert inhoudelijke resultaten op!



Slimme werkwijze

..met praktisch materiaal

waar het hele veld gebruik van kan maken!



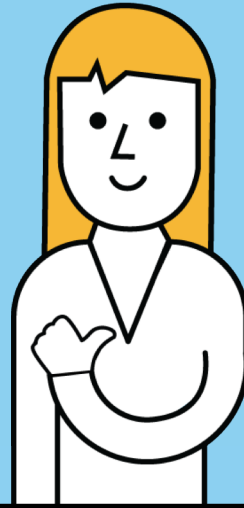
Blik in de toekomst

..van datagedreven aanpak in het MBO!

Plannen voor de toekomst

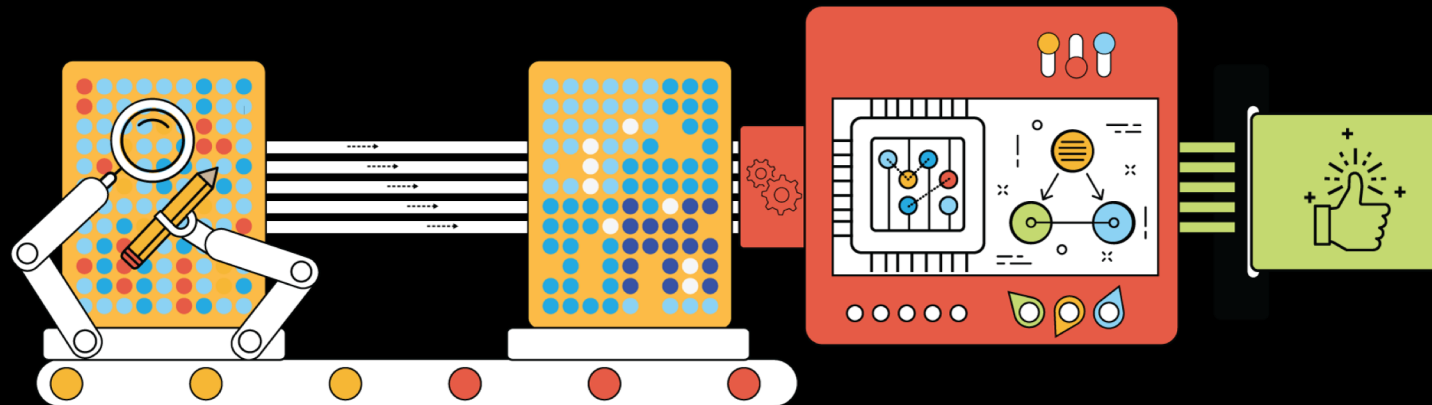
Plan

1. Groei in aantal deelnemende instellingen in de Datacoalitie
2. Samen bepalen we de thema's, focus en aanpak
3. Samen aan de slag: onderzoeken en projecten uitvoeren
4. Delen van kennis en inzichten i.s.m. Programma DOO



Verwachtingenmodel uitval mbo niveau 2

Uitval voorkomen door studenten nóg beter begeleiden



Probleemstelling

- Uitval van studenten is een probleem bij de meeste mbo instellingen. Studenten die uitvallen hebben een significant kleiner kans op werk en redden zich minder goed in de maatschappij. De focus is gelegd op niveau 2 omdat het aandeel studenten zonder startkwalificatie daar het grootst is.
- Verklaring van uitval vaak pas achteraf duidelijk. Noodzaak om vooraf een inschatting te kunnen maken van uitval om alle studenten optimaal te begeleiden naar een diploma. Een passende verdeling van middelen kan daarbij helpen.
- **1. Welke factoren zijn er bij DUO bekend over studenten die instromen op niveau 2 en kunnen deze factoren gebruikt worden om een accurate voorspelling van uitval te doen?**
- **2. Wat is de behoefte van onderwijsteams met betrekking tot de inzet van een verwachtingenmodel en hoe zouden onderwijsteams het willen toepassen?**

Methode

1. Ontwikkeling verwachtingenmodel door Martin Belder (data-scientist bij DUO)

- a) Netwerkanalyse dat gebruik maakt van gradient-boosting techniek.
- b) Niveau 2 studenten gelabeld als “uitval” en “geen uitval”. Model getraind met historische data van duo vanaf cohort 2015 t/m 201
- c) Model bevat per student de volgende data (zie volgende slide voor details)
 - persoonsgegevens,
 - onderwijsverleden,
 - Omgevingskenmerken
- d) Focus op mbo niveau 2 studenten die nieuw instromen in het mbo

2. Interview potentiële gebruikers van het verwachtingenmodel

- opleidingsmanagers
- directieleden (onderwijs)
- coaches

Resultaten

Model

Voorspelling van de meest waarschijnlijke uitkomst na 1 jaar studeren

Hoe presteert het model?

Nauwkeurigheid 83%, AUC score 0,81.

Dit betekent dat het model de uitval van 4 van de 5 studenten die uitvallen kan voorspellen

Dashboard

Dashboard in omgeving van DUO

- Gebaseerd op kenmerken van studenten die instromen in het mbo
- Voorspelde % studenten met risico op uitval op team niveau
- Belangrijkste kenmerken opgenomen die verband hebben met voorspelde % op uitval
- Mogelijkheid om data te exporteren en verder te gebruiken binnen de instelling

Voorspelde uitval in het MBO niveau 2

Instelling

ROC van Amsterdam

Locatie

Alle vestigingen

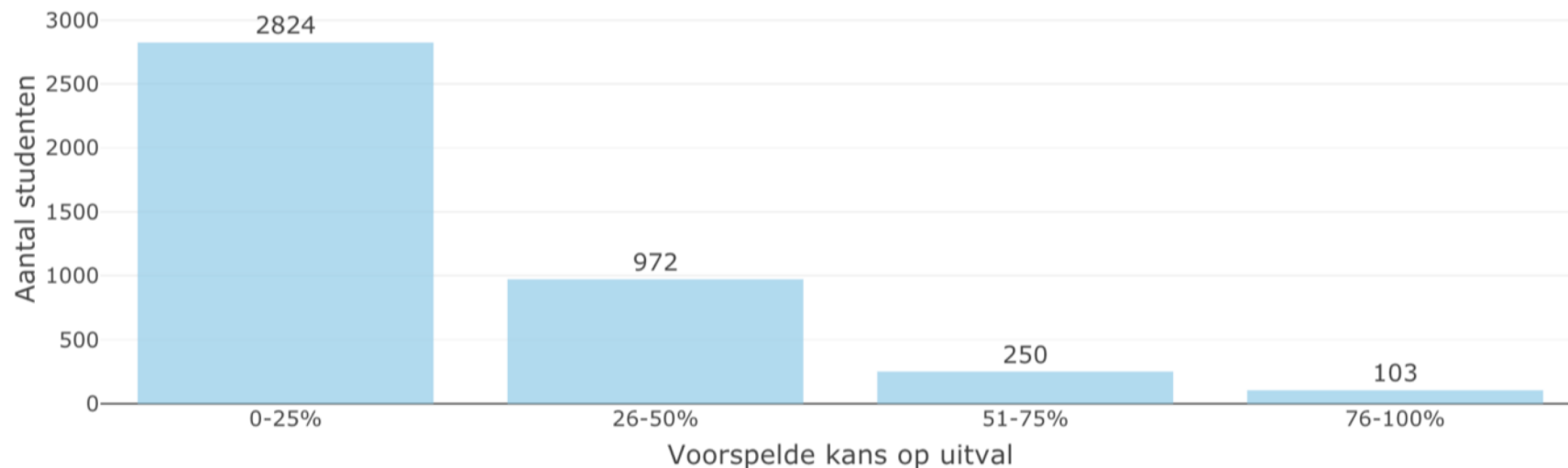
Team

GEEN SELECTIE

Voorspelling

Kenmerken

Toelichting voorspelmodel



CSV

Excel

PDF

Instelling	Locatie	Team	Aantal studenten	Voorspelde risicogroep
ROC van Amsterdam	MBO College Airport		301	123 leerlingen (41%)
ROC van Amsterdam	MBO College Centrum		197	49 leerlingen (25%)
ROC van Amsterdam	MBO College Hilversum		675	114 leerlingen (17%)

Een voorspelling van risicogroepen op basis van instroomeigenschappen:
Waar zitten grote groepen studenten met een hoger risico / kans op uitval?

Casus: gesprek over onderwijskwaliteit

Werkelijke uitval in te vullen door instelling

Exported data					
Instelling	Locatie	TEAM	Voorspelling in risicogroep	# studenten risicogroep	Werkelijke uitval *
ROC van Amsterdam	MBO College Centrum	A	16%	7	13%
ROC van Amsterdam	MBO College Centrum	B	15%	4	16%
ROC van Amsterdam	MBO College Centrum	C	30%	27	36%
ROC van Amsterdam	MBO College Centrum	D	27%	10	22%
TOTALEN			22%	48	

* Werkelijke uitval betreft in dit voorbeeld fictieve cijfers

Toepassing van het verwachtingenmodel

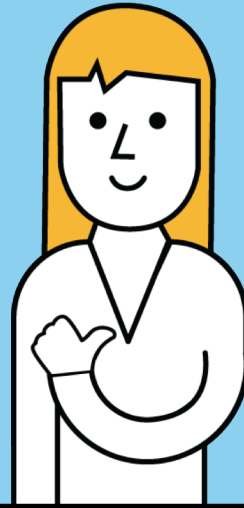
- Interviews:
 - Positief over de toepassing van het model
 - Behoeftte aan voorspelling van uitval op groepsniveau als het kan zelfs per student
 - Inzicht in doorstroom van studenten naar niveau 3 ook als waardevolle verbetering gezien
 - Switchers ook in de populatie opnemen.
 - Ook uitval conform definitie jaarresultaat en VSV is interessant
 - Uitsplitsing van de resultaten naar opleidingsniveau en uitbreiding niveau 3 en 4
 - Tijdsreeksopnemen in dashboard

Conclusie

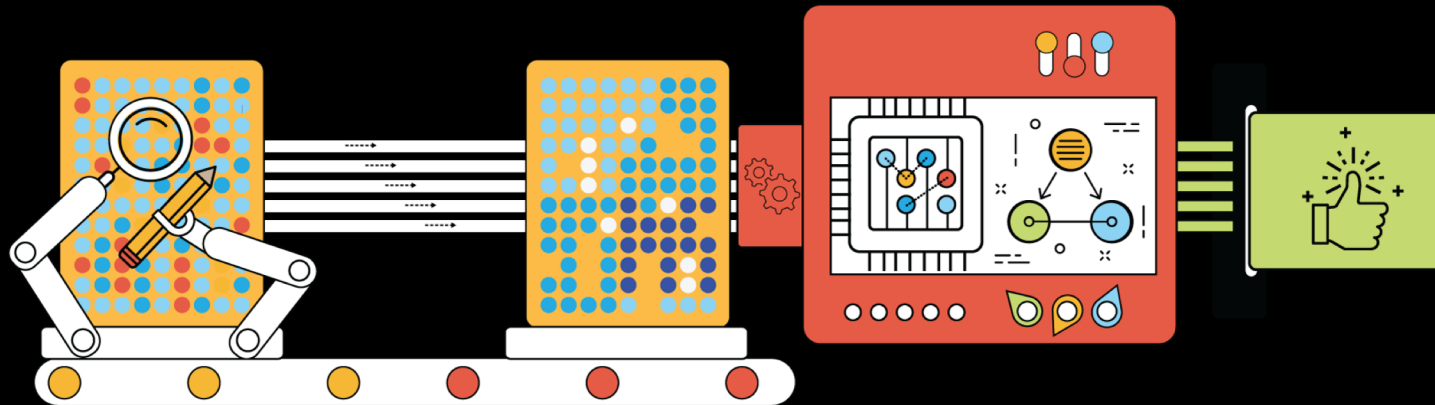
- **Met behulp van de bij duo bekende data is het mogelijk om met een accuraatheid van ~80% uitval van niveau 2 studenten te voorspellen.**
 - Het model is niet in staat om de kans op doorstroom naar mbo niveau 3 en diplomeren te onderscheiden.
- **Opleidingsmanager zijn enthousiast over de potentie van het model**
 - opleidingsmanagers zijn ook geïnteresseerd in de voorspelde kans op doorstroom naar mbo niveau 3.
 - Onderzoek naar mogelijke verdieping en verbreding van het model is noodzakelijk.
- **Dashboard gerealiseerd met de verwachte uitval per instelling, locatie en opleidingsteam**

Vervolg (aug21 – jan22)

- Onderzoek naar de implementatievraagstukken van het model in het onderwijs.
- Randvoorwaardelijke zaken onderzoeken (tijdigheid data instroom, DPIA, IT faciliteiten DUO)
- De mogelijkheden verkennen voor verdere verdieping en verbreding van het verwachtingenmodel (bijvoorbeeld niveau 3 & 4 toevoegen of data uit andere bronnen)



Integraal Teamdashboard en onderzoek naar datakwaliteit



WG1: integraal teamdashboard

Beoogd doel

- Meer inzicht verkrijgen in de behoefte van onderwijsteams (tbv welke werkprocessen, welke besluitvorming, op welk moment en in welke vorm)
- Meer inzicht verkrijgen in welke indicatoren en data onderwijsinstellingen reeds aan onderwijsteams beschikbaar stellen

Verderliggend doel:

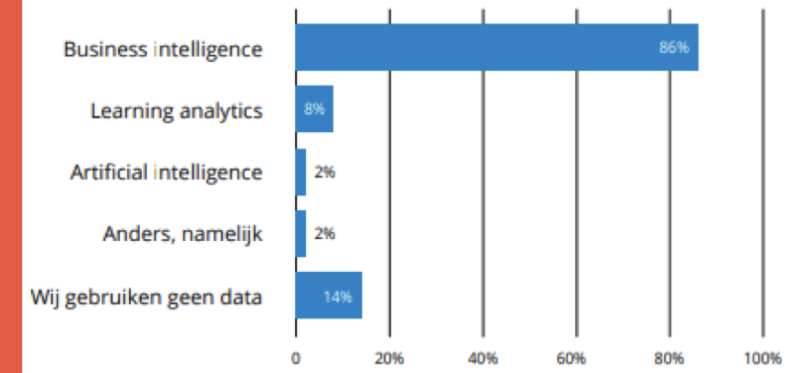
- Meer ervaring opdoen met het bedenken en ontwerpen van praktische, in teams toepasbare, tools en vormen van datagedreven werken/-besluitvorming

WG1: integraal teamdashboard

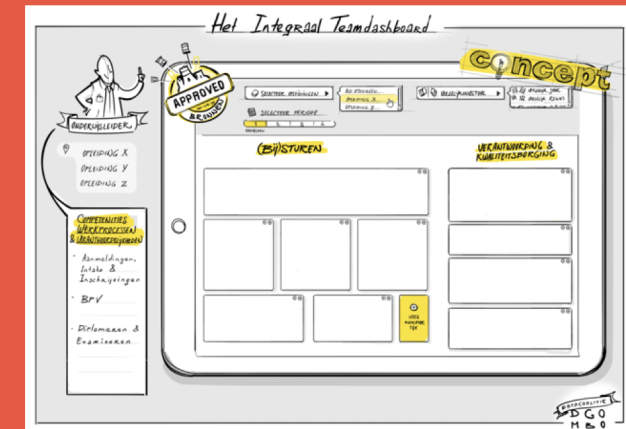
Aanpak

- Deskresearch: inventarisatie en analyse van bestaande indicatoren en dataleveringen aan onderwijsteams
- Interviews met onderwijsleiders, docenten, SLB'ers en kwaliteit zorgmedewerkers a.d.h.v praatplaten: inspiratie en behoefte inventarisatie
- Ontwerp: uitwerking van inzichten tot 'eerste functionele beschrijving' van Integraal Teamdashboard incl praatplaten

Figuur 1.11 Waar gebruiken de instellingen data voor?



ICT monitor 2019 (Kennisnet, 2020)



Resultaten

- *Onderzoeksfase:*
 - Elke instelling kent zijn eigen applicatielandschap en databronnen
 - Er is een gemeenschappelijke referentiearchitectuur
 - Instellingen hebben overeenkomende rollen en processen
 - Gebruikers hebben overeenkomende informatie-/databehoefte
 - Inzicht gekregen in de gewenste indicatoren per doelgroep
 - Instellingen maken (nog te) beperkt gebruik van de beschikbare data
 - Doelgroepen zien toegevoegde waarde van 'HIT'
- *Ontwerpfase:*
 - Eerste functioneel ontwerp voor HIT
 - Vier praatplaten; doelgroep specifiek

WG1: integraal teamdashboard

Advies & voorstel:

- Functioneel concept is een denkrichting
 - Per instelling een integraal teamdashboard?
 - Een integraal teamdashboard visualiseert de beschikbare data binnen een instelling
 - Maar wat is de kwaliteit van deze beschikbare data?
-
- **Advies werkgroep aan stuurgroep:** focus op onderzoek naar datakwaliteit / -integriteit t.b.v. integraal teamdashboard.

WG1:Stand van zaken 'Onderzoek datakwaliteit'

Opdracht:

- Doe onderzoek naar datakwaliteit in relatie tot het Integraal Teamdashboard
- Werk naar een gedeelde definitie van datakwaliteit
- Het resultaat van het onderzoek moet instelling- en sturingsonafhankelijk zijn
- Het resultaat moet iedere instelling vooruit helpen op het gebied van (het verbeteren van) datakwaliteit.

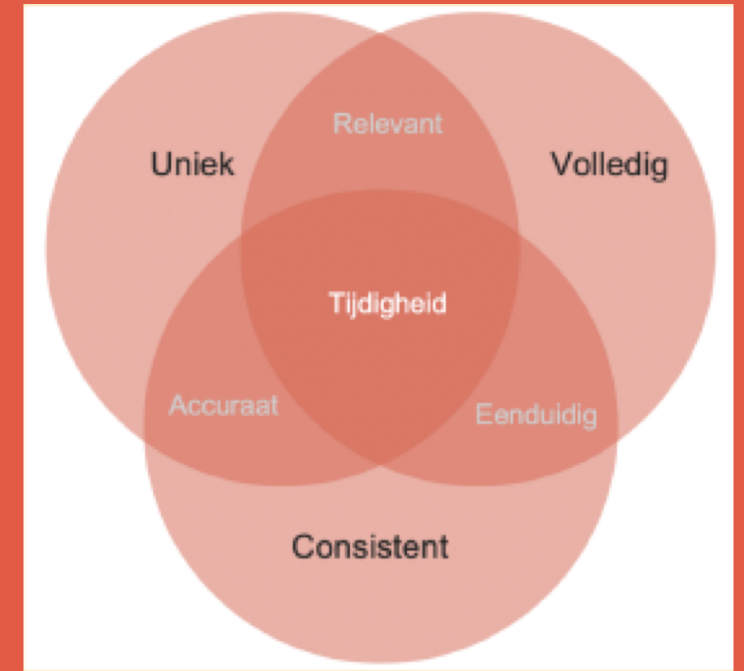
Beoogd doel:

- Gezamenlijk meer inzicht verkrijgen in definitie en de praktijk van datakwaliteit in het mbo
- Praktische handvatten bieden waarmee onderwijsinstellingen datakwaliteit kunnen onderzoeken, meten, monitoren en verbeteren

WG1:Stand van zaken 'Onderzoek datakwaliteit'

Aanpak (1)

- Deskresearch: literatuurstudie, actuele definitie en handvaten
- Concept instrument/stappenplan ontwerpen voor:
 - Diagnose: de datakwaliteit in beeld kunnen brengen aan de hand van een raamwerk
 - Oorzaak: kunnen achterhalen waardoor die kwaliteit onbekend of onvoldoende was
 - Interventie: acties kunnen ondernemen om de kwaliteit op orde te krijgen?



Aanpak (2)

- Inventariseren relevante actuele casussen van deelnemende instellingen waarin datakwaliteit een aandachtspunt:
 - Casus 1: De kwaliteit & registratie van exitgesprekken
 - Casus 2: De kwaliteit van aan-en afwezigheidsregistratie
 - Casus 3: De kwaliteit van onderwijslocaties & aanbieders in RIO
- Casus onderzoek a.d.h.v. het ontwikkelde instrument
- Aanscherpen onderzoeksinstrument datakwaliteit

WG1:Stand van zaken 'Onderzoek datakwaliteit'

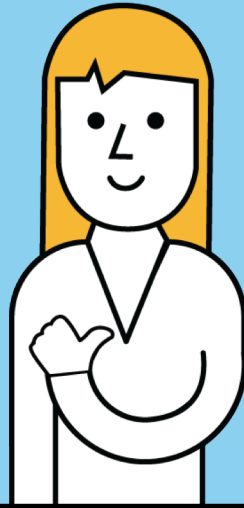
Raamwerk datakwaliteit		
Onderdeel	Groep	Vragen
Vorbereiding & inventarisatie	Aanleiding	4
	Betrokkenen	5
Situatie in beeld	De scope	4
	De indicator	3
	De maker	4
	De gebruiker	4
	Uniek	4
Diagnose en onderliggende oorzaak	Relevant	2
	Volledig	4
	Eenduidig	2
	Consistent	3
	Accuraat	2
	Tijdigheid	3
	Validiteit	1
	Stel de prioriteiten vast	2
Interventie	Onderzoek de oorzaak	2
	Ontwerp een verbetering	4
	Monitor het effect en evalueer	2

Stand van zaken (tussen)resultaten

- Literatuurstudie is afgerond
- Het concept onderzoeksinstrument is gerealiseerd
- Interviews met betrokkenen zijn gehouden

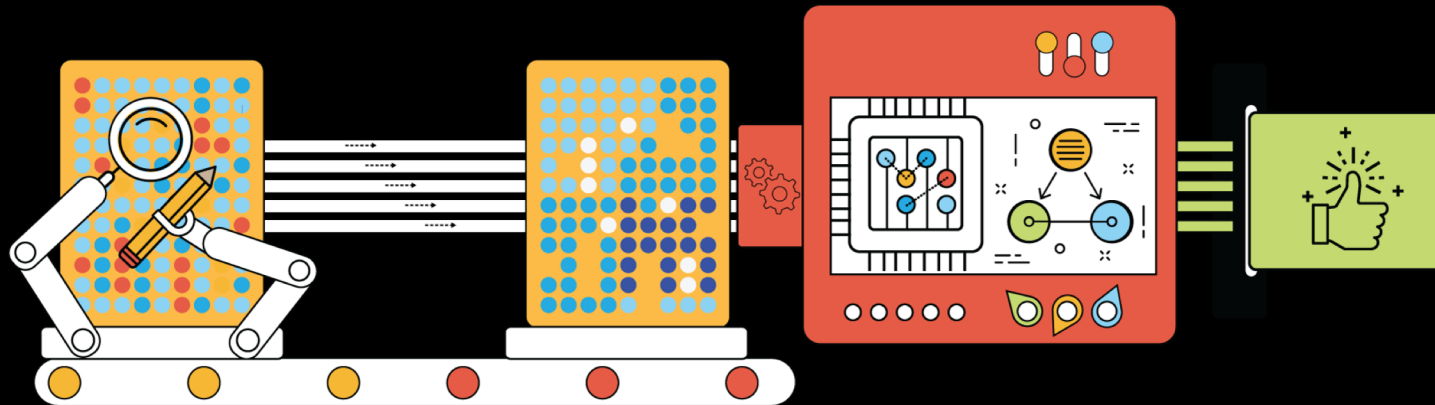
Vervolgstappen:

- Onderzoeksinstrument aanscherpen met opgedane ervaringen
- Onderzoeksinstrument visualiseren in aansprekende infographic
- Begeleidende notitie bij onderzoeksinstrument uitwerken



Verbeteren databeschikbaarheid

Werving, Instroom en Doorstroom



WG2: data over werving, instroom & doorstroom

Beoogd doel

Hoofddoel:

- **Betere aansluiting van dataleveringen op de behoefte van de instellingen**

Subdoel:

- Inzicht krijgen in huidig data-aanbod versus de databehoeft van de instellingen
- Meer inzicht in waar studenten vandaan komen (instroom) en waar ze heen gaan na de opleiding (uitstroom)

Gekozen aanpak en proces

1. Focus op databestanden en niet op diensten (dashboards)
2. Informatiebehoefte van instellingen in beeld brengen
3. Relevante databestanden doorlopen en beoordelen
4. Samen analyseren: waar sluiten vraag en aanbod niet aan?
5. Verbeterwensen geformuleerd en afgestemd met DUO, OCW en MBO Raad
6. Activiteitenplan opgesteld hoe te komen tot verbeterde aansluiting

Conceptplanning dataleveringen



betrokkenen



Dienst Uitvoering Onderwijs
Ministerie van Onderwijs, Cultuur en
Wetenschap



data coalitie
DGO

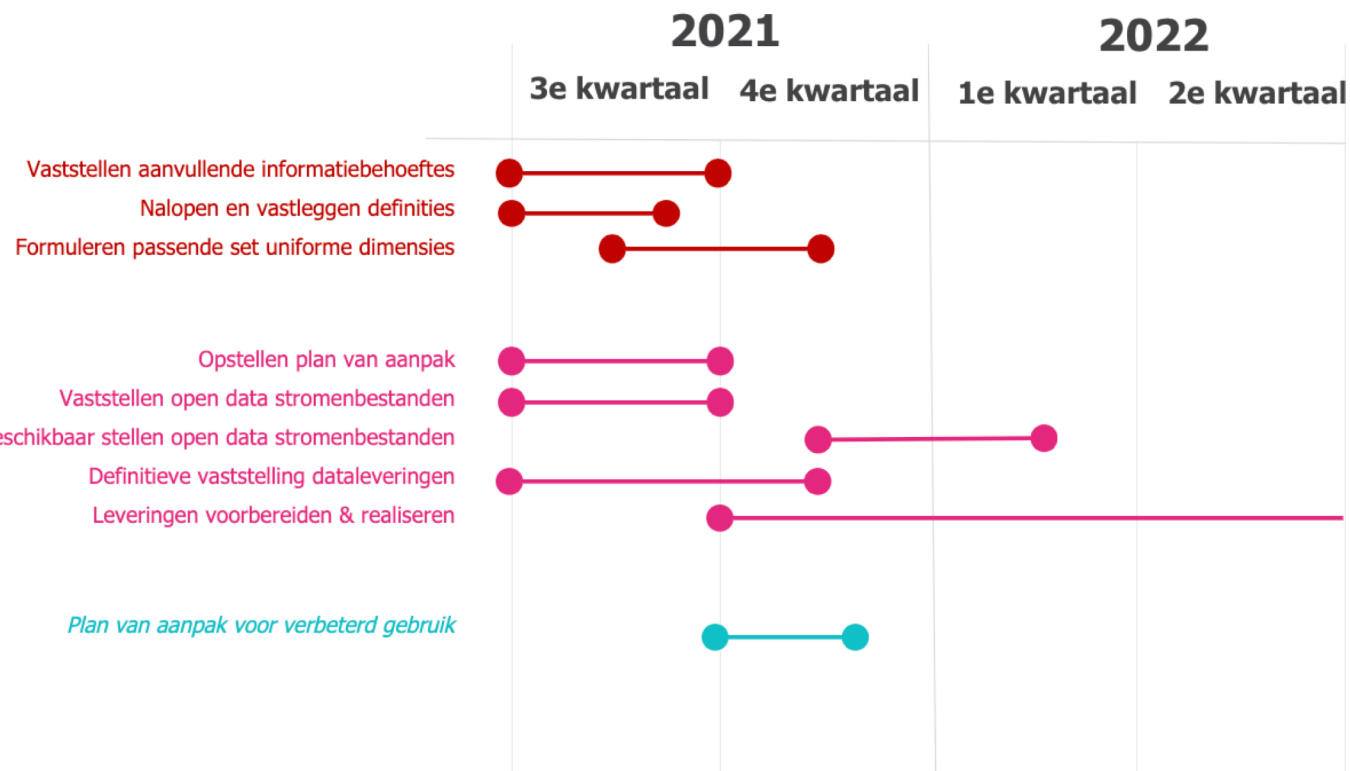


MBO Raad

Vertrekpunt
vaststellen en
toetsen

Dataleveringen
op orde

Bekendheid &
bruikbaarheid
verbeteren



Waar zijn we nu mee bezig?

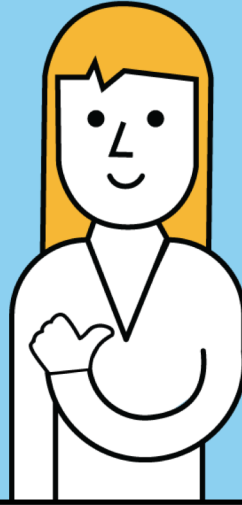
- We bekijken de informatiebehoefte aan de hand van 6 thema's
 - Macrodoelmatigheid
 - Instroom in het mbo
 - Student op de juiste plek *
 - Doorstroom binnen het MBO
 - Doorstroom naar het HBO
 - Uitstroom naar arbeid *
- We hanteren de volgende uitgangspunten voor datalevering
 - Open data waar mogelijk, anders achter de PIN, maar altijd conform AVG
 - Dimensies worden redundant geleverd (bovenliggende niveaus leveren we mee net als naam + code waar mogelijk)
 - beschikbaar via CSV-bestand en een API
 - We bouwen op naar 5 jaar historische data (starten met 2 jaar wegens beschikbaarheid RIO)
 - We hanteren altijd de laatste stand van zaken binnen de dimensies. Bijvoorbeeld bij gemeentelijke indeling of wijzigingen van kwalificaties.

*Voor de thema's uitstroom naar arbeid & student op de juiste plek kijken we naar een samenwerking met CBS

Wanneer komen de bestanden?

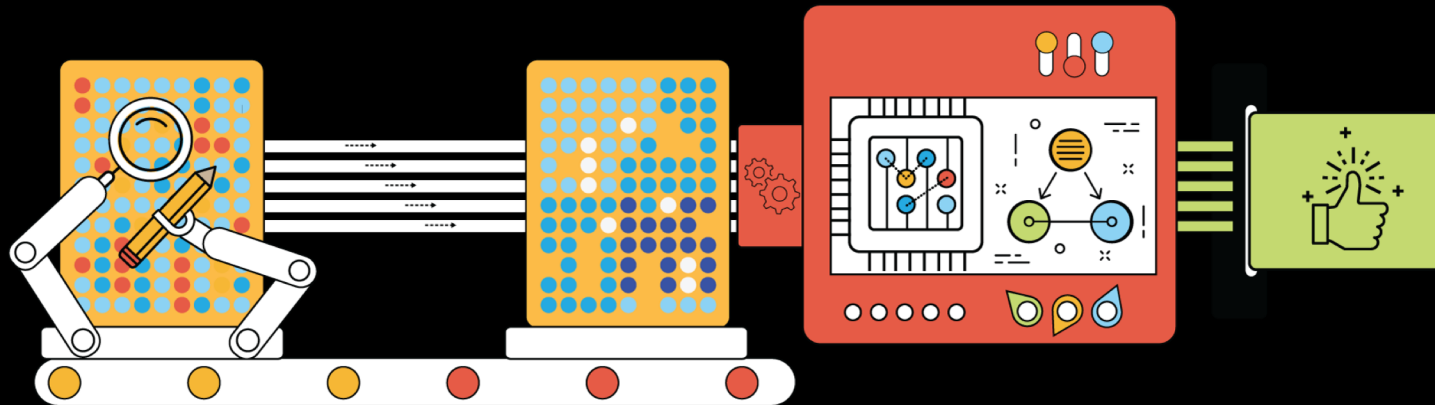
Planning

- In januari willen we voor de thema's instroom in het MBO en doorstroom naar het HBO de bestanden afronden met de werkgroep DGO.
- In juni 2022 willen de thema's doorstroom binnen het MBO en macrodoelmatigheid afronden met de werkgroep DGO.
- Voor de thema's uitstroom naar arbeid en student op juiste plek zijn we (deels) afhankelijk van data van CBS en/of SBB. Het streven is om deze thema's ook in 2022 te behandelen en te ontsluiten.



Datacoalitie Datagedreven onderzoek

Studenten nóg beter begeleiden met een datagedreven aanpak!



Doe je mee?

Wat vragen we van je

- Inhoudelijke bijdrage
 - vertegenwoordiging per instelling, die in meerdere werkgroepen kunnen deelnemen
 - indien nodig ondersteuning van collega's (bijv onderwijs & kwaliteit, marketing, informatiemanager, dataspecialist, FG'er)
- Afvaardiging voor de stuurgroep (vanuit bestuur of management)
- Inschatting uren inzet: 2 tot 4u per week gedurende projectperiode, afhankelijk van rol/afspraken
- Financiële bijdrage obv begroting: 10-15k (projectbegeleiding, analyse kennis/capaciteit)

Doe je mee?

Wat levert het je op?

- Opdoen van kennis en praktijkervaring met datagedreven werken/onderzoeken
- Beantwoording van onderzoeksvragen en thema's met data
- Toegang tot externe data om te integreren met instellingsdata
- Kijkje achter de schermen bij DUO (CBS, SBB, gemeente, etc..)
- Actief netwerk van experts op gebied datagedreven werken in het mbo

Vervolg

Alle resultaten staan op: www.datagedrevenonderzoekmbo.nl

Wil je meer weten over de datacoalitie en de toekomstplannen?

of

Wil je aanmelden voor de datacoalitie?

tom@hutspot.nl en we nemen contact op.

Contact

Tom – tom@hutspot.nl

Paulo – pmoekotte@rocvantwente.nl

Harry – h.vdevis@albeda.nl

Roel – r.klaver@rocva.nl

Willem-Jan – willemjan@newnodes.nl