



Pilothub Studiedata en AI

Centre for Educational Data Analytics (CEDA)



Kennisbank › Doorstroom en aansluiting tussen mbo en hbo

Doorstroom en aansluiting tussen mbo en hbo

Gepubliceerd: 18 maart 2022

Studenten met een mbo-4 diploma kunnen doorstromen naar hbo. Dit kan mogelijk even vanzelfsprekend.

Kabinet wil loten bij universiteiten en hogescholen weer toestaan

Numerus Fixus Uit onderzoek van de onderwijsinspectie bleek dat studenten c

De 'juiste' geneeskundestudent blijkt vaak wit, vrouw en kind van hoogopgeleiden – door de selectiemethode

HvA en VU intensiveren samenwerking om iedere student het juiste leerpad te bieden

Mbo'ers verdienen meer dan hbo'ers en vooral in deze branches



'Juist bij flexibiliteit heeft student behoefte aan structuur'

Studentenaantallen dalen sterker dan eerder voorspeld, bezuinigingen dreigen

Kabinet onderzoekt dwingendere opleidingskeuzes

[Nieuws](#) | [de redactie](#)

29 mei 2024 | De demografische krimp en de arbeidsmarkttekortingen zetten de vrije studiekeuze van jongeren onder druk. In de toekomst kan de overheid de capaciteit op bepaalde opleidingen begrenzen en studenten nog actiever richting opleidingen voor tekortsectoren sturen. Binnenkort komt OCW met een onderzoek naar sturende maatregelen.

De student staat centraal?!

Kamer geeft hoger onderwijs opnieuw het recht om buitenlandse studenten te weren



'Bsa past niet bij modulair onderwijs'

Mogelijkheden met data analyse



How Data Collaboration Platforms Can Help Companies Build Better AI

Onderzoek | Artificial intelligence | mbo

Kan AI studie-uitval in het mbo voorspellen?

12 maart 2024

Aanbesteding voor pan-Europese inkoop van clouddiensten
gepubliceerd: OCRE 2024

De universiteit heeft nu haar eigen glazen bol: nieuw model voorspelt toekomstige instroom van studenten

29 apr 2024 Mathijs Noij



Sinds de komst van kunstmatige intelligentie en deep learning wordt Moore's Law verpulverd

A new frontier: Composable data systems

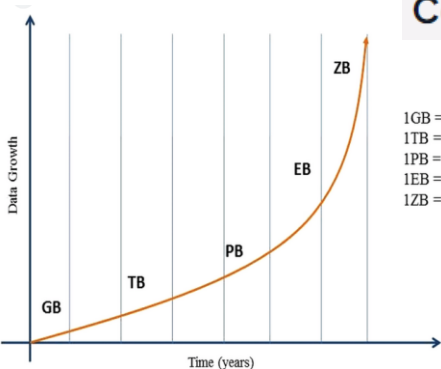
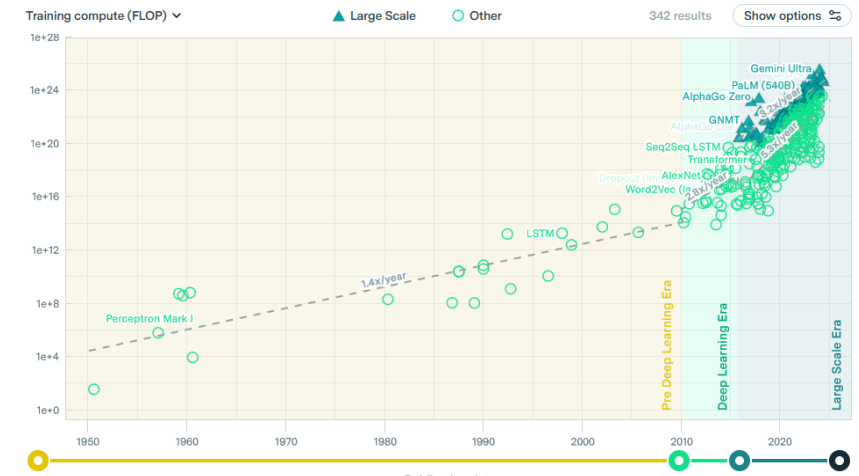
29 november 2023

Open Onderwijs API wint de Gouden API Award

APRIL 24, 2024

Empowering students through Data-Driven Learning

Training Compute of Notable Machine Learning Systems Over Time



1GB = 10⁹ bytes
1TB = 10¹² bytes
1PB = 10¹⁵ bytes
1EB = 10¹⁸ bytes
1ZB = 10²¹ bytes

The rise of data towards 'big data' over the years

Ontwikkelingen

- Groeiende behoefte aan **snelheid** om tot inzichten te komen
- Informatiebehoefte steeds in **verandering** (regionale krimp, capaciteitsbekostiging, kansengelijkheid, welzijn, doorstroom, LLO)
- Steeds **meer (uniforme) data**
- Cloud- en machine learning **technieken steeds toegankelijker**
- Behoeftte aan **digitale autonomie**

Huidige situatie

Vaak starre
standaardrapportages

Data op een interne server

Beperkte capaciteit (enkelingen
met ervaring en gevoel voor de
data)...

...die >80% van de tijd bezig
met ontsluiten data

Soms vender lockin

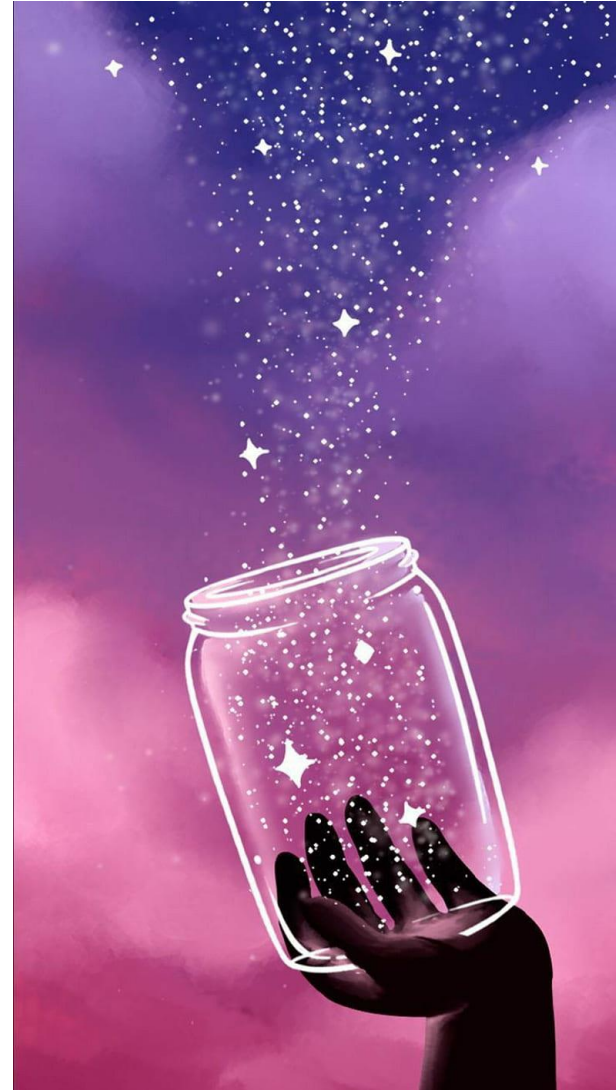
Oplossingen lastig deel- en
schaalbaar (instellingen vinden
zelf wiel uit)

Maken we optimaal gebruik van
de ontwikkeling op technisch
vlak?



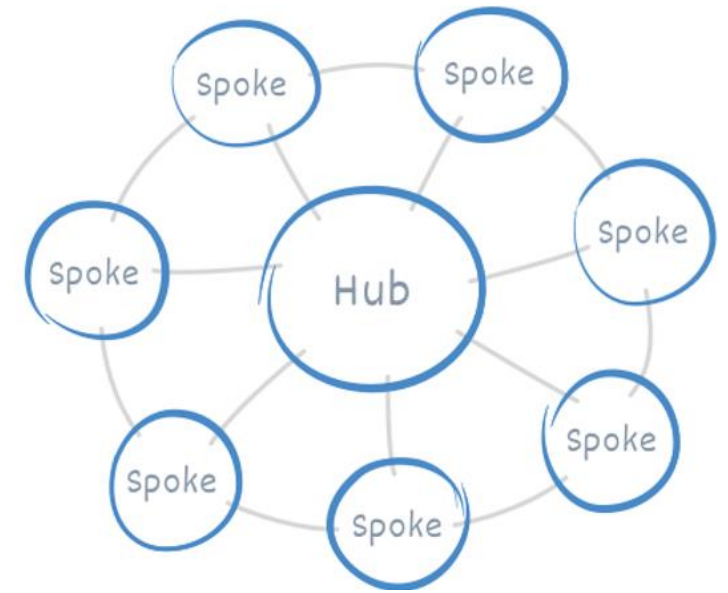
Wat als we...

- ... gezamenlijke standaarden en 'ways of working' laten ontstaan
- ... samenwerken aan gedeelde analytics oplossingen
- ... die oplossingen deelbaar maken (analytics in a box)



Uitgangspunten CEDA

- “*Most of the smartest people work somewhere else*”
- Voor en door het onderwijs:
 - Modulair (bouwstenen)
 - Open source (aanpasbaar en deelbaar)
 - Kennisdeling (community en documentatie)
- Doelen CEDA
 - hogere kwaliteit van inzichten
 - meer expertise
 - snellere ‘return on information’
 - instelling blijft in control



Eerste stappen

- Collega's van instellingen zetten best practices om naar deelbaar product
 - Instroomvoorspelling
 - Wisselstroom package en script
 - Inschrijvingen dashboard
 - ICHO package
- Ways of working
 - Hoe zorgen we dat echt het deelbaar is?
 - **Daarvoor hebben we jullie nodig!**



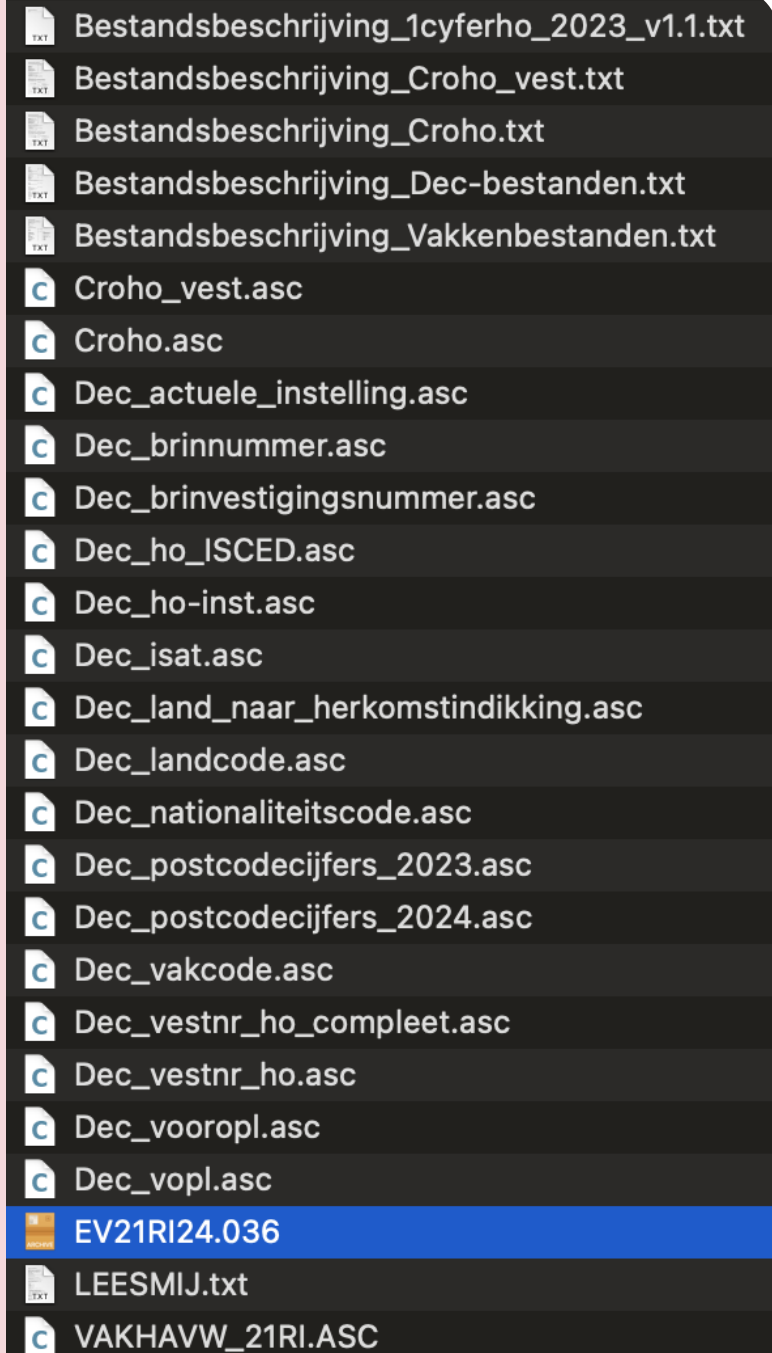
1 cijfer

5 July 2024



Wat is het 1 cijfer?

- Extract uit instellingssysteem
- Aangevuld door DUO en CBS
- Basis voor bekostiging vanuit DUO
- Uniform voor hbo en wo (en misschien zelfs mbo)
- Mede bedoeld voor kwaliteitszorg
- Rijke bron



Waarom een package?

- Het is een rijke bron
- Het is een exotisch bestandsformaat
- Jaren ervaring nodig om het te doorgronden
- Nog nooit gedaan...





Wat bedoel je met ‘rijke bron’?

- Inschrijvingshistorie van instelling
- Postcodeindeling naar gemeente per jaar
- ISCED-indeling per opleiding
- Croho naar CrohoActueel
- Eindexamencijfers
- Historie in het hoger onderwijs
- Internationale studenten

Wat is dan het package geworden?

- Een **CLI** dat is te installeren als **Python**-package
- Installeren op de voor Python standaard manier
- ***pip install eencijfer***
- **Configureerbaar** op de command line
- **Open source** op Github (<https://github.com/ennnngb/eencijfer>)

```
(eencijfer-py3.9) (base) bramenning@Brams-MacBook-Pro ruw % eencijfer convert --help
```

```
Usage: eencijfer convert [OPTIONS]
```

```
Convert eencijfer-files to desired exportformat, with or without PII.
```

```
Options
```

```
--source-dir
```

```
PATH
```

```
--result-dir
```

```
PATH
```

```
--export-format
```

```
[csv|parquet|xlsx|duckdb]
```

```
--use-column-converters
```

```
-c
```

```
--not-use-column-converters
```

```
-C
```

```
--remove-pii
```

```
-p
```

```
--do-not-remove-pii
```

```
-P
```

```
--add-local-id
```

```
-s
```

```
--do-not-add-local-id
```

```
-S
```

```
--help
```

```
(eencijfer-py3.9) (base) bramenning@Brams-MacBook-Pro ruw %
```

eencijfer convert

Optie	Beschrijving
--help	Beschrijving van alle opties die bij een commando horen
--export-format	omzetten naar csv, excel, parquet of duckdb
--use-column-converters	formaat van kolommen (datum, integer, categorien) bij exporteren omzetten.
--remove-pii	verwijderen van kolommen als Burgerservicenummer, Onderwijsnummer en pseudonimiseren van kolom Persoonsgebondennummer
--add-local-id	toevoegen van lokaal nummer (bijv. Studentnummer)

eencijfer create-assets

Omzetten naar analyse-sets.
Let op: werkt alleen met parquet-bestanden als input!

Optie	Beschrijving
--help	Beschrijving van alle opties die bij een commando horen
--export-format	omzetten naar csv, excel, parquet of duckdb

datasets	Beschrijving
eencijfer	Het EV-bestand aangevuld met bijv. schoolnamen, vooropleiding, lokale naam van opleiding, isced-indeling, croho-sector, et cetera.
cohorten	Per regel een student die voor het eerst instroom in de instelling, indicatoren (uitval, switch, rendement) zoals die in de Prestatieafspraken werden gebruikt.
¹ eindexamencijfers	Per student, vooropleiding, diplomajaar en vak een regel met de school- en eindexamenresultaten.



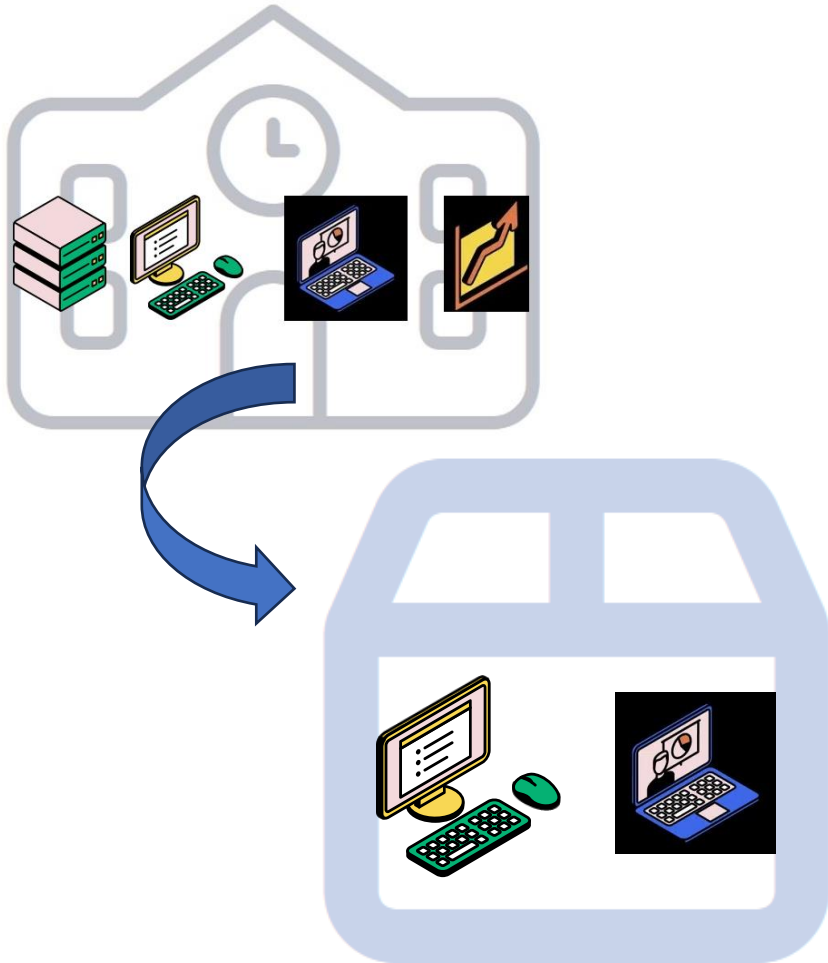
Ik wil gewoon tabellen

5 July 2024



Hoe lost het CEDA dit op?

Best practices uit vooroplopende instellingen verzamelen en deze technisch toegankelijk maken.



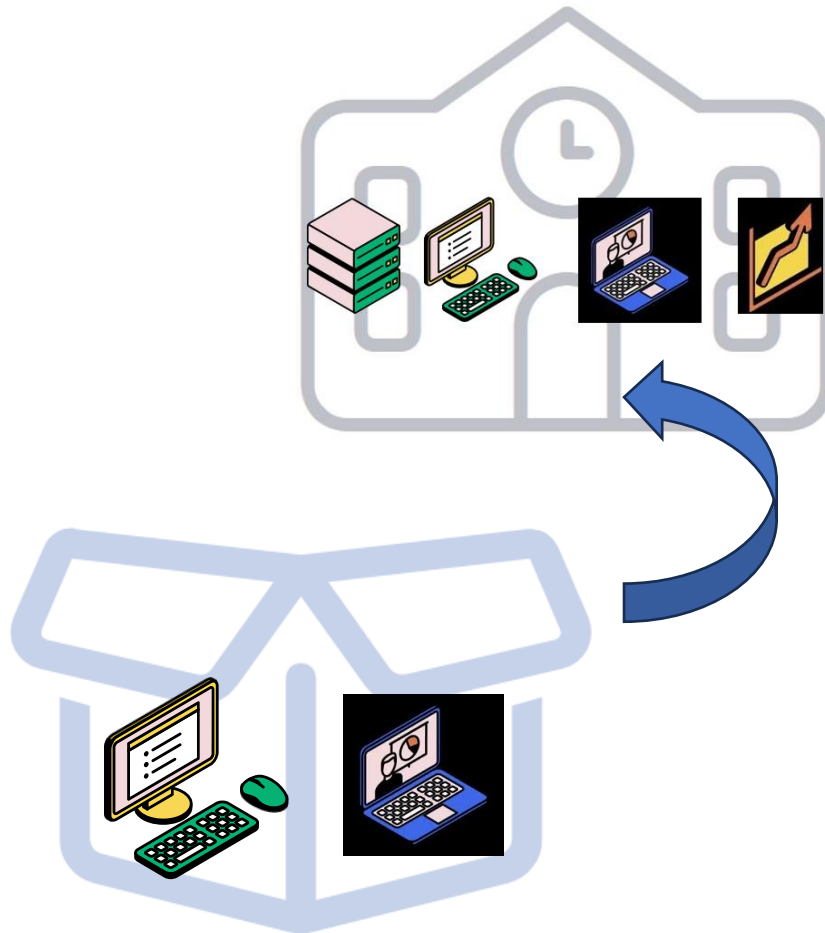
Analytics



Analytics-in-a-box

Q1 – Q2 2024

Hoe lost het CEDA dit op?



Deze analytics-in-a-box delen met andere instellingen en praktisch toegankelijk maken.

Analytics-in-a-box

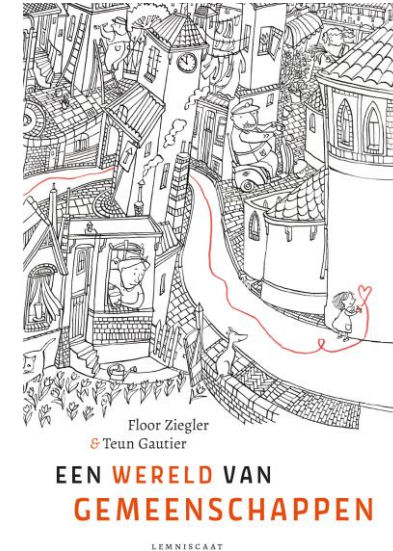


Analytics-out-of-the-box

Q3 – Q4 2024

Next steps

- Volgende workshop
 - 24 juni
 - Wisselstroom in R
 - 1CijferHO PowerBI dashboard
- Volgende serie PoCs
 - Call open augustus / september
- Strategische sessie
 - ~2e week van september
- Afrondende praktische sessies
 - ~ Ergens na de zomer
 - [DAIR Seminar](#)
 - OWD 2024



*"Ask not what your community
can do for you..."*



