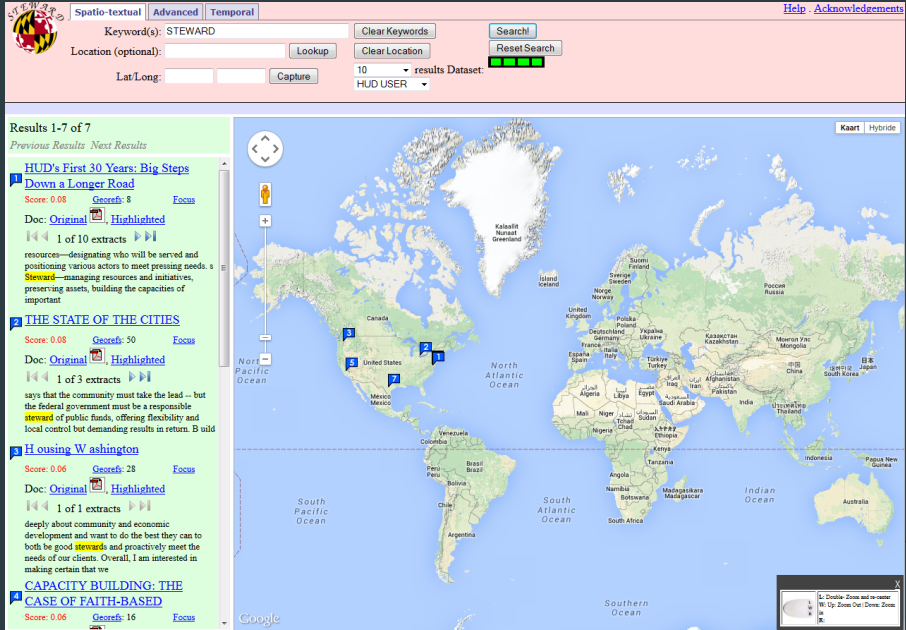


# Grafische geografisch- gebaseerde selectie van nieuws

Thomas van Ophem

Begeleider: Raphael Poss

► STEWARD



# Gerelateerd onderzoek

## ► STEWARD

- Extraheren, opvragen en visualiseren referenties geografische locaties in tekst
- Nieuws, monitoring ziektes etc.

The screenshot displays the STEWARD web application interface. At the top, there are tabs for 'Spatio-textual', 'Advanced', and 'Temporal'. Below these, a search bar contains the keyword 'STEWARD'. To the right of the search bar are buttons for 'Clear Keywords', 'Search!', 'Clear Location', and 'Reset Search'. Below the search bar, there are fields for 'Location (optional):', 'Lat/Long:', and 'Capture'. A 'LookUp' button is also present. On the right side, there is a 'results Dataset:' dropdown menu showing '10' and 'HUD USER'. Below the search bar, there are links for 'Help' and 'Acknowledgements'.

The main content area is divided into two panels. The left panel, titled 'Results 1-7 of 7', lists search results. The first result is 'HUD's First 30 Years: Big Steps Down a Longer Road' with a score of 0.08 and 8 geo-references. The second result is 'THE STATE OF THE CITIES' with a score of 0.08 and 50 geo-references. The third result is 'Housing Washington' with a score of 0.06 and 28 geo-references. The right panel shows a world map with markers indicating the locations of the documents. The map is labeled 'Kaart | Hybrid' and includes a Google logo at the bottom.

# Gerelateerd onderzoek

## ► STEWARD

- Extraheren, opvragen en visualiseren referenties geografische locaties in tekst
- Nieuws, monitoring ziektes etc.
- Nog geen gebied selecteren

The screenshot displays the STEWARD web application interface. At the top, there are tabs for 'Spatio-textual', 'Advanced', and 'Temporal'. Below these, a search bar contains the keyword 'STEWARD'. To the right of the search bar are buttons for 'Clear Keywords', 'Search!', 'Clear Location', and 'Reset Search'. Below the search bar, there is a 'Location (optional):' field with a 'Lookup' button, and a 'Lat Long:' field with a 'Capture' button. A 'results Dataset:' dropdown menu is set to 'HUD USER'. On the left side, under 'Results 1-7 of 7', there are three search results listed. Each result includes a title, a score, a 'Geared' value, a 'Focus' button, and a 'Doc: Original' link. The first result is 'HUD's First 30 Years: Big Steps Down a Longer Road' with a score of 0.08 and 8 geared items. The second result is 'THE STATE OF THE CITIES' with a score of 0.08 and 50 geared items. The third result is 'Housing Washington' with a score of 0.06 and 28 geared items. On the right side, there is a world map with a search bar and a 'Kaart' button. The map shows various countries and regions, with some areas highlighted in green. A legend in the bottom right corner of the map indicates that green areas represent 'Geared' regions and red areas represent 'Not Geared' regions.

# Gerelateerd onderzoek

## ► NewsStand



# Gerelateerd onderzoek

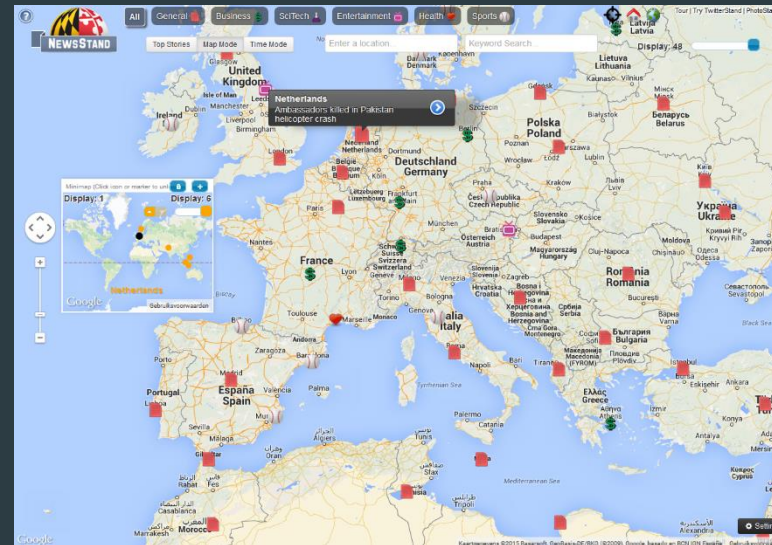
- ▶ NewsStand
  - ▶ Inzichtelijk maken geografische informatie nieuws berichten





# Gerelateerd onderzoek

- NewsStand
  - Inzichtelijk maken geografische informatie nieuws berichten
  - Op kaart weergegeven waar bericht over gaat
  - Nog geen gebied selecteren



# Ontwerp keuzes

- ▶ Web applicatie of standalone?



# Ontwerp keuzes

- ▶ Web applicatie of standalone?
- ▶ Client side/Server side

# Ontwerp keuzes

- ▶ Web applicatie of standalone?
- ▶ Client side/Server side
- ▶ Python/CherryPy/Cheetah
- ▶ SQLite

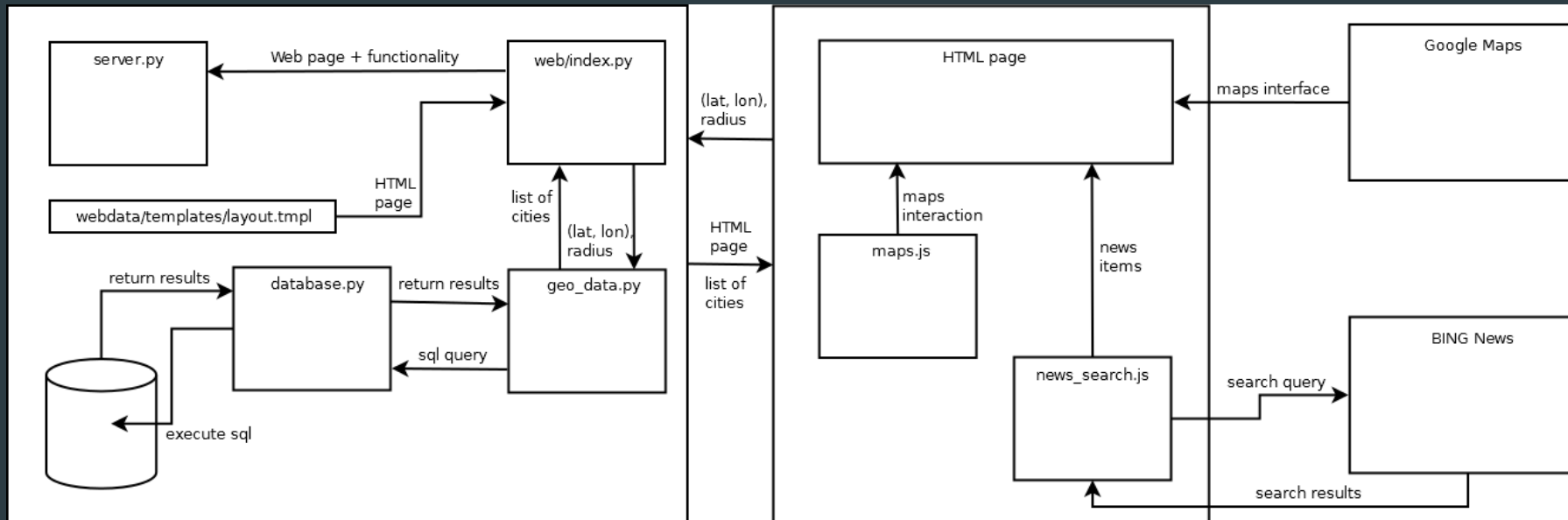
# Ontwerp keuzes

- ▶ Web applicatie of standalone?
- ▶ Client side/Server side
- ▶ Python/CherryPy/Cheetah
- ▶ SQLite
- ▶ Geodata → [geonames.org](http://geonames.org)
- ▶ Nieuwsbronnen → BING News

# Ontwerp keuzes

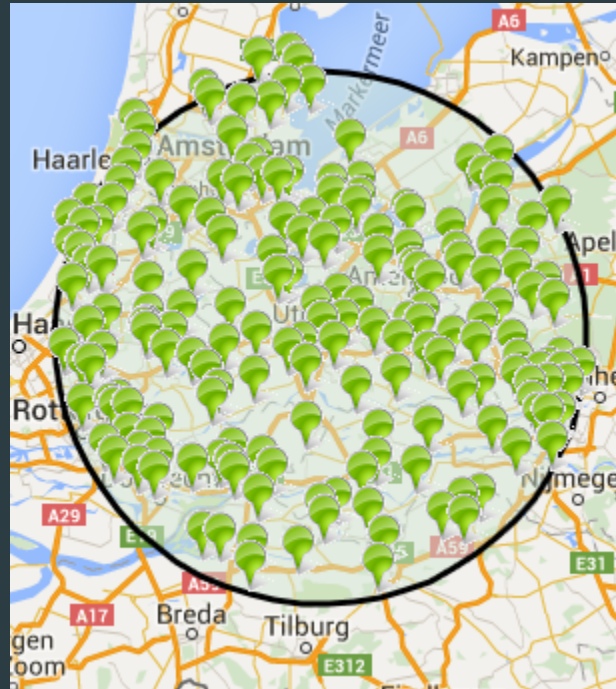
- ▶ Web applicatie of standalone?
- ▶ Client side/Server side
- ▶ Python/CherryPy/Cheetah
- ▶ SQLite
- ▶ Geodata → [geonames.org](http://geonames.org)
- ▶ Nieuwsbronnen → BING News
- ▶ Ideaal ontwerp?

# Architectuur



# Gebiedsrepresentatie

## ► Waarom?



# Gebiedsrepresentatie

- ▶ Waarom?
- ▶ Hoe?





# Gebiedsrepresentatie

- ▶ Waarom?
- ▶ Hoe?
- ▶ Willekeurig? Grootste steden? Hoofdsteden?



# Gebiedsrepresentatie

- ▶ Twee algoritmes
- ▶ 15 steden

# Gebiedsrepresentatie

- ▶ Twee algoritmes
- ▶ 15 steden
- ▶ Grootste stad zoeken

Split():

**Input:** list of cities

**Result:** Reduced list of cities representing the area

```
while number of result < number of cities to represent the area do
    find biggest city;
    add to result;

    foreach city do
        calculate bearing;
        add to corresponding sublist (nw/ne/se/sw);
    end

    foreach sublist (nw/ne/se/sw) do
        if number of cities ≤ 5 then
            find biggest city;
            add to result;
        else
            Split(sublist);
            add to result;
        end
    end
end
end
```

# Gebiedsrepresentatie

- ▶ Twee algoritmes
- ▶ 15 steden
- ▶ Grootste stad zoeken
- ▶ Splitsen in vier deel gebieden

Split():

**Input:** list of cities

**Result:** Reduced list of cities representing the area

```
while number of result < number of cities to represent the area do
    find biggest city;
    add to result;

    foreach city do
        calculate bearing;
        add to corresponding sublist (nw/ne/se/sw);
    end

    foreach sublist (nw/ne/se/sw) do
        if number of cities <= 5 then
            find biggest city;
            add to result;
        else
            Split(sublist);
            add to result;
        end
    end
end
end
```

# Gebiedsrepresentatie

- ▶ Twee algoritmes
- ▶ 15 steden
- ▶ Grootste stad zoeken
- ▶ Splitsen in vier deel gebieden
- ▶ Elk deel gebied opnieuw

Split():

**Input:** list of cities

**Result:** Reduced list of cities representing the area

```
while number of result < number of cities to represent the area do
    find biggest city;
    add to result;

    foreach city do
        calculate bearing;
        add to corresponding sublist (nw/ne/se/sw);
    end

    foreach sublist (nw/ne/se/sw) do
        if number of cities <= 5 then
            find biggest city;
            add to result;
        else
            Split(sublist);
            add to result;
        end
    end
end
end
```

# Demonstratie

[Home](#) [Contact](#) [F.A.Q.](#)

Met deze applicatie kunt u nieuws zoeken door een gebied te selecteren op de kaart. Dit selecteren doet u door een punt te kiezen op de kaart, de linker muisknop ingedrukt te houden en te slepen tot het gebied groot genoeg is. Zodra u de linker muisknop los laat wordt er naar nieuws gezocht. Het is helaas niet mogelijk om de kaart te slepen, als u naar een andere gebied wilt kunt u dit doen door in en uit te zoomen. In en uit zoomen kan door op de kaart naar boven en beneden te scrollen.



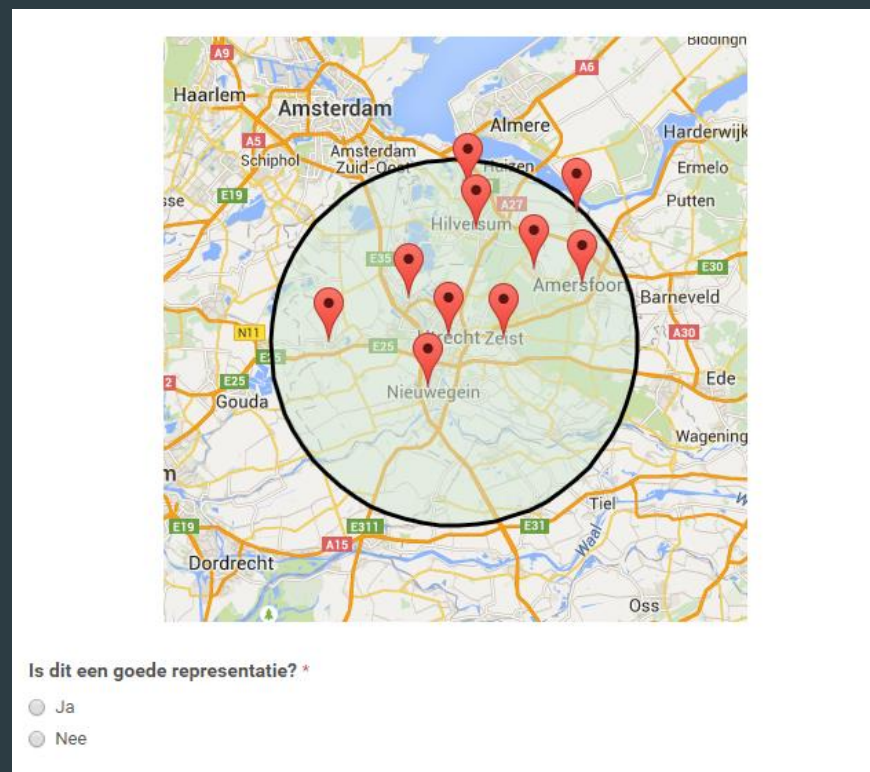
[Westfriesgasthuis in Hoornt wint gastvrijheid award](#)  
[Gewonde bij explosie op jacht in haven Hoornt](#)  
[Hoornt eert Team Brunel tijdens Volvo Ocean Race](#)  
[Hoornt - Man opgepakt na uitschelden agent](#)  
[Heroische reis naar Kaap Hoornt](#)

News for Amstelveen...

[Sportverenigingsmarkt 2015 in Stadshart Amstelveen](#)  
[Opinie: 'Betaald parkeren in bos zorgt voor meer parkeerdruk in Amstelveen'](#)  
[Amstelveen in top-3 beste woonplaats](#)  
[Sportmarkt op het Stadsplein](#)  
[Vluchtelingenwerk in groter huis](#)  
[Sesamstraat-poes Dikkie Dik weer theater in Stadshart Amstelveen is 'Beste Winkelcentrum van Nederland'](#)  
['Mijn vader is vermoord met 82 steken' \[video\]](#)  
[Hoofddorp - Twee aanhoudingen na mishandeling op snelweg](#)  
[KLM werkt aan gezonde financiële basis](#)

# Onderzoek gebiedsrepresentatie

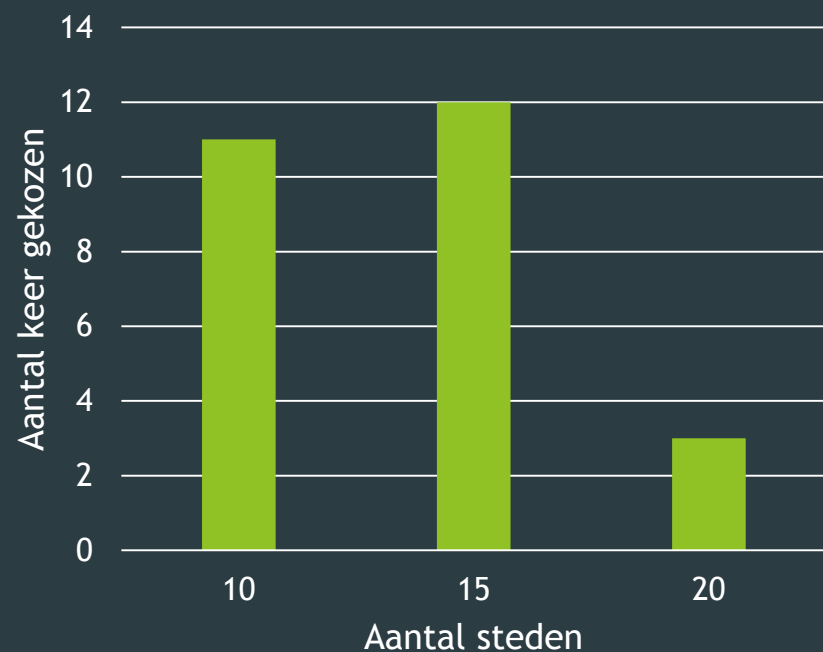
- ▶ Vier gebieden, straal 25km, 100km, 500km en 1000km
- ▶ Voor elk gebied keuze uit 10, 15 en 20 steden
- ▶ Vergelijking tussen de twee algoritmes



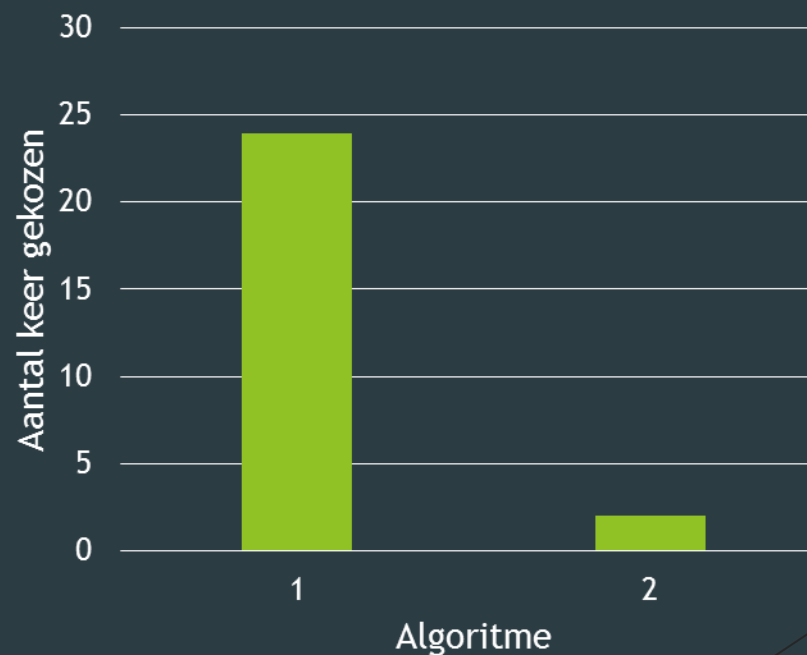


# Resultaten - onderzoek gebiedsrepresentatie - 25km

## Benodigde aantal steden

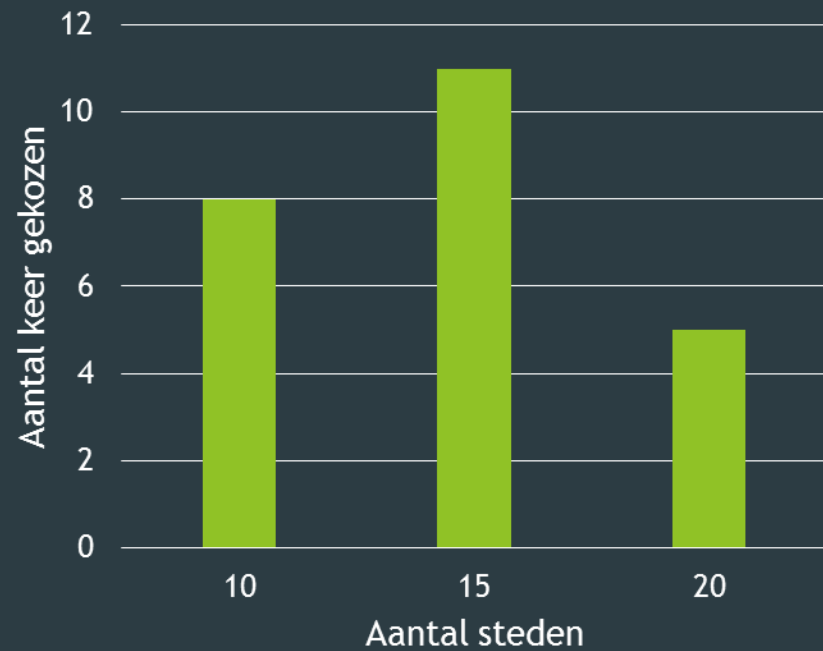


## Algoritme keuze

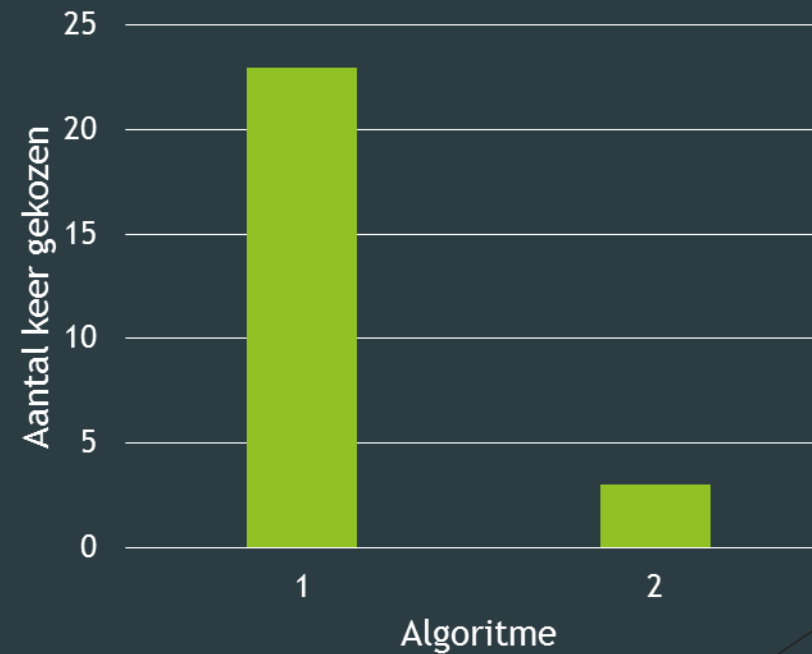


# Resultaten - onderzoek gebiedsrepresentatie - 100km

## Benodigde aantal steden

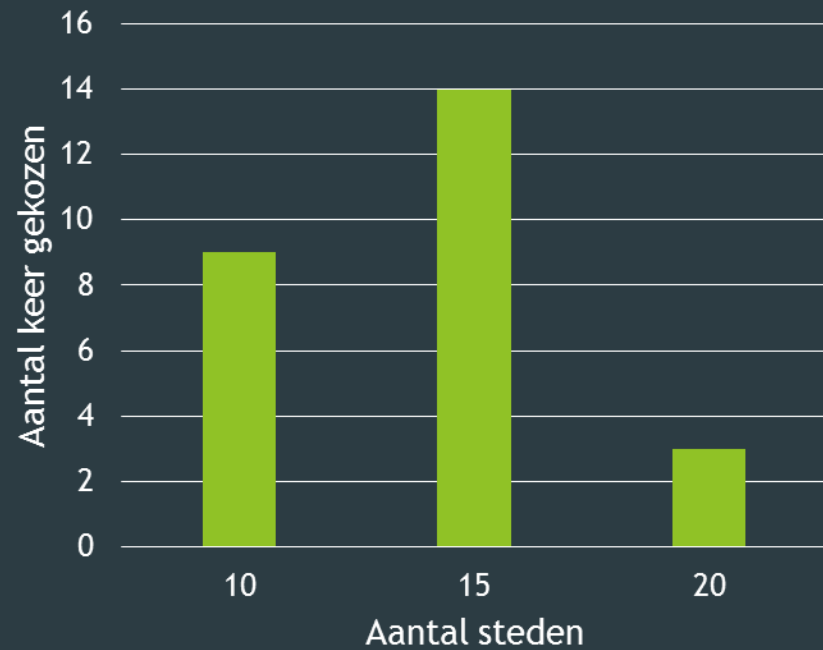


## Algoritme keuze

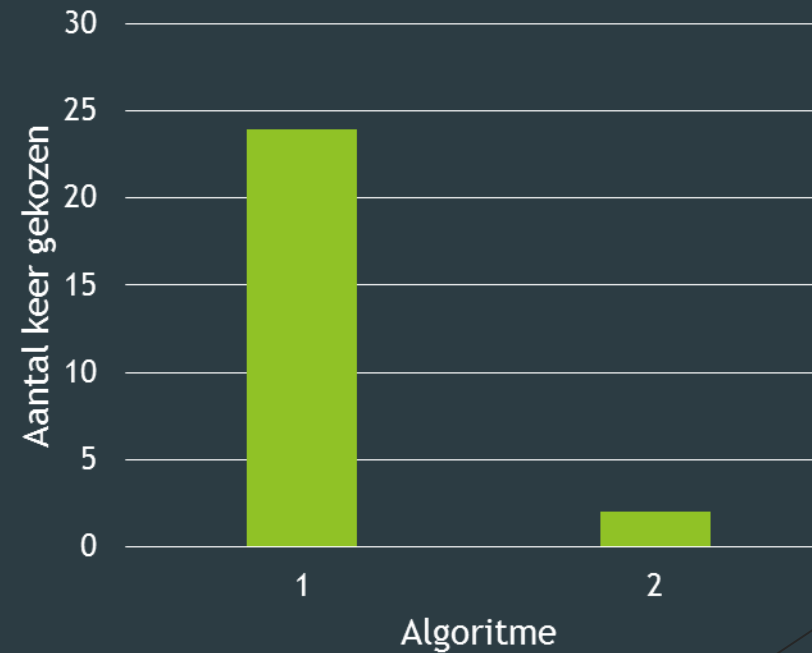


# Resultaten - onderzoek gebiedsrepresentatie - 500km

## Benodigde aantal steden

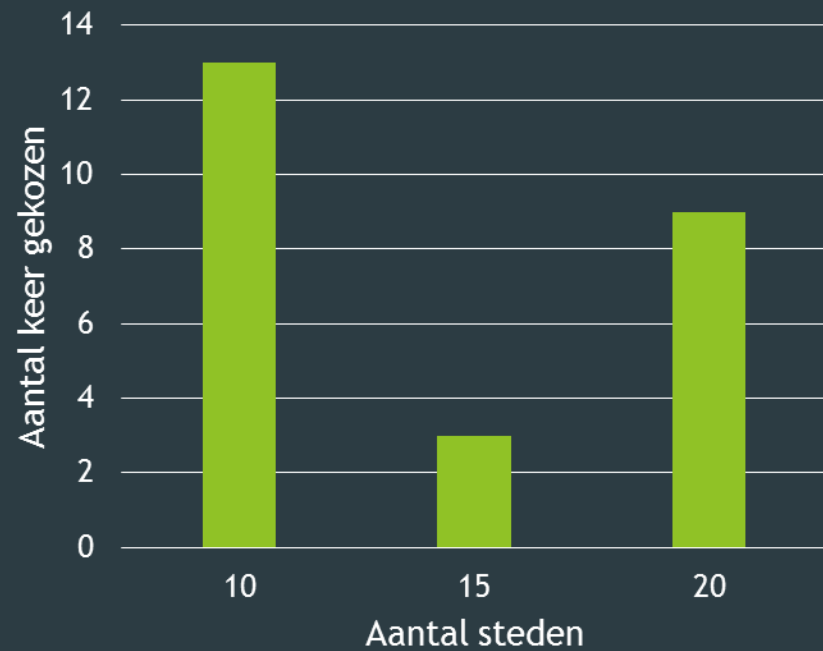


## Algoritme keuze

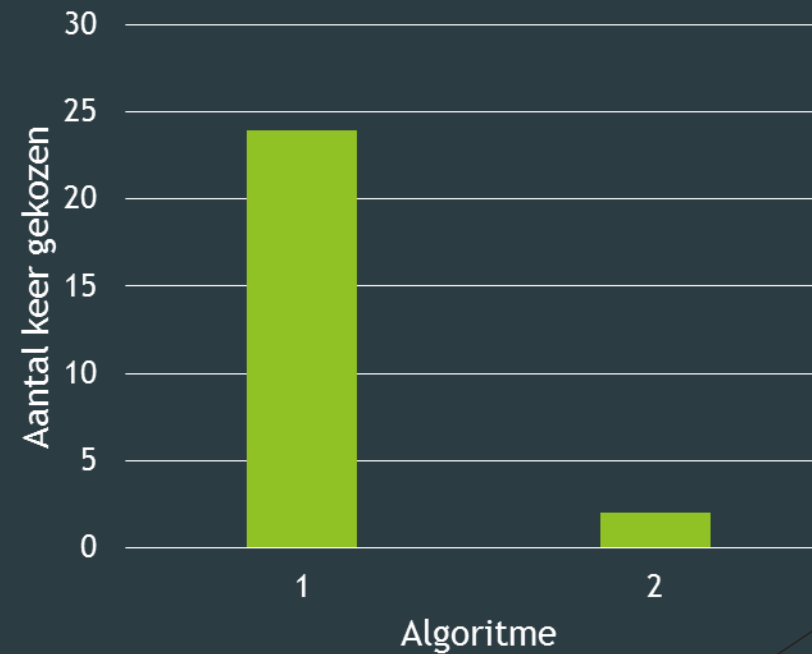


# Resultaten - onderzoek gebiedsrepresentatie - 1000km

## Benodigde aantal steden



## Algoritme keuze



# Onderzoek toegevoegde waarde

- ▶ Zoek zoveel mogelijk interessant nieuws uit Noord Holland

# Onderzoek toegevoegde waarde

- ▶ Zoek zoveel mogelijk interessant nieuws uit Noord Holland
- ▶ Eerst 5 minuten op “normale” manier
- ▶ Daarna 5 minuten met ontwikkelde applicatie

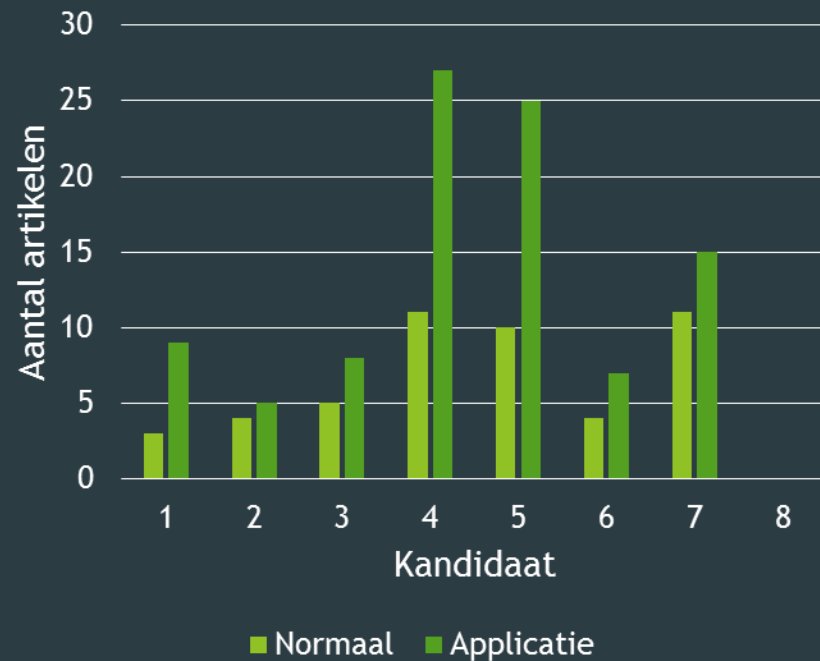
# Onderzoek toegevoegde waarde

- ▶ Zoek zoveel mogelijk interessant nieuws uit Noord Holland
- ▶ Eerst 5 minuten op “normale” manier
- ▶ Daarna 5 minuten met ontwikkelde applicatie
- ▶ Hoeveel nieuws gevonden?
- ▶ Heeft de applicatie toegevoegde waarde en waarom?

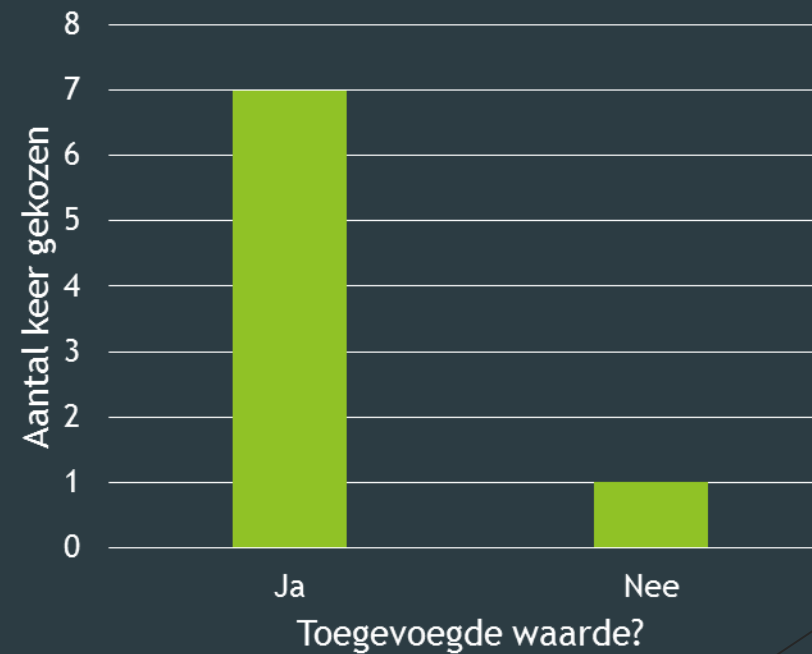


# Resultaten - onderzoek toegevoegde waarde

## Aantal gevonden artikelen



## Toegevoegde waarde



# Conclusie

- ▶ Toegevoegde waarde? Ja

# Conclusie

- ▶ Toegevoegde waarde? Ja
- ▶ Extra onderzoek gebiedsrepresentatie

# Conclusie

- ▶ Toegevoegde waarde? Ja
- ▶ Extra onderzoek gebiedsrepresentatie
- ▶ Hopelijk in de toekomst beschikbaar
- ▶ Toevoegingen: onderwerp en tijd

# Conclusie

- ▶ Toegevoegde waarde? Ja
  - ▶ Extra onderzoek gebiedsrepresentatie
  - ▶ Hopelijk in de toekomst beschikbaar
  - ▶ Toevoegingen: onderwerp en tijd
- 
- ▶ Vragen?