



# Met de computer naar advertenties kijken

---

Melvin Wevers

November 20, 2017

Digital Humanities Group - KNAW Humanities Cluster

# Inhoudsopgave

1. Advertenties als historische bron
2. Metadata
3. Tekst
4. Beeld
5. SIAMESE
6. Volgende stappen

## **Advertenties als historische bron**

---

# Achtergrond

- “insight into the ideals and aspirations of past realities. They show the state of technology, the social functions of products, and provide information on the society in which a product was sold” (Marchand, 1985)
- Bijna 20 miljoen krantenadvertenties (1890-1990)
- Hoe kunnen we computers inzetten om deze grote hoeveelheid advertenties te bestuderen?

# Advertentie

- Metadata
- Tekst
- Afbeelding



Figure 1: *Limburger Koerier*, April 13, 1938

# Metadata

---

# Metadatavelden

- Krantentitel
- Datum
- Grootte
- Positie op pagina
- Paginanummer
- Totaal aantal pagina's in krant

# Hoe groot waren de advertenties?

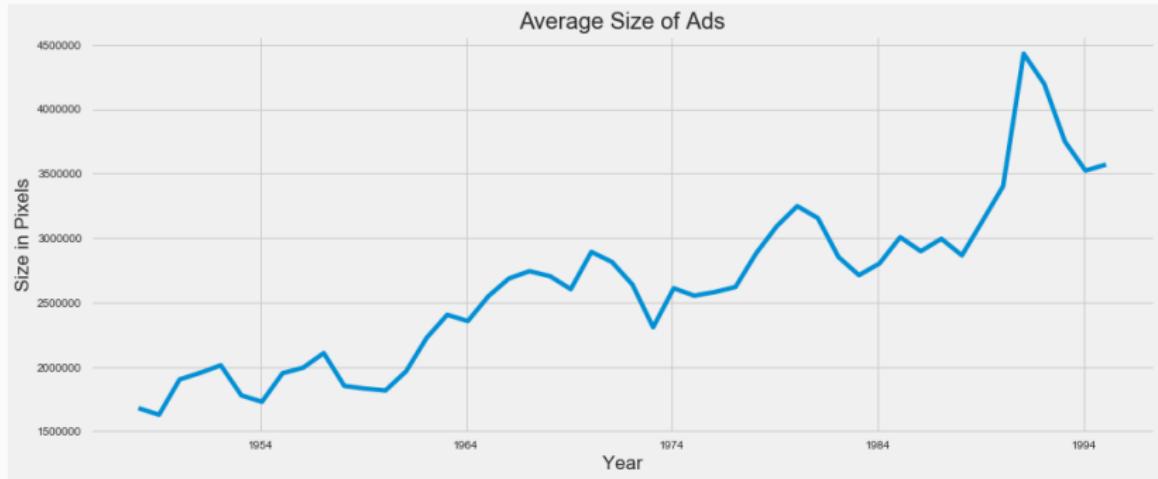
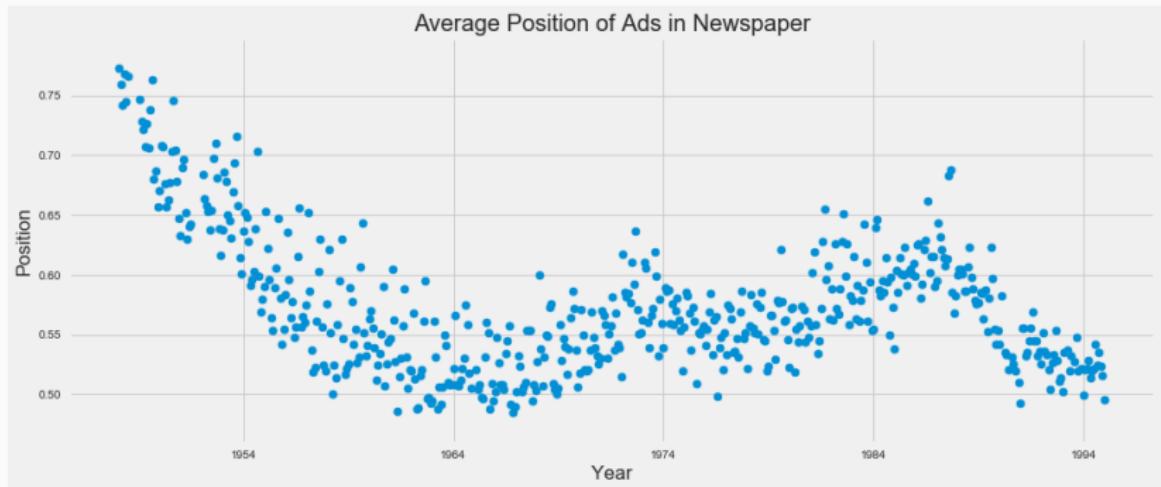


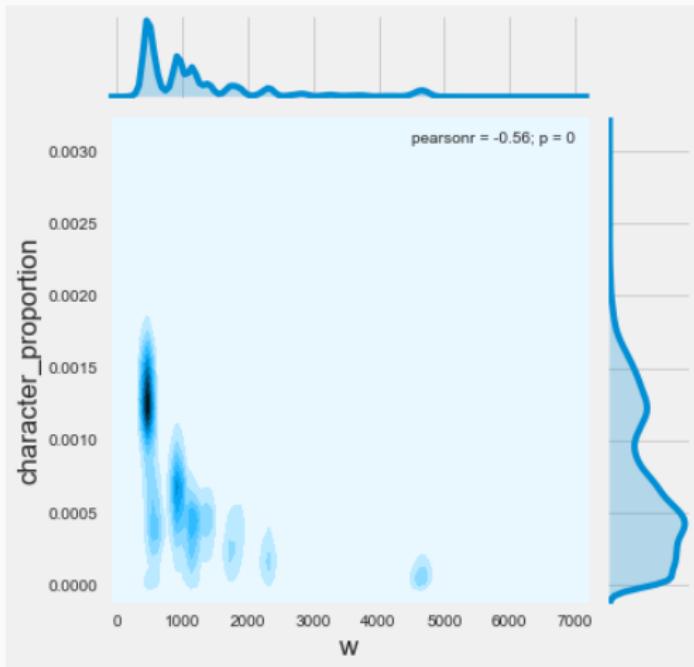
Figure 2: Gemiddelde grootte van advertenties in AH en NRC

# Waar stonden de advertenties in de krant?



**Figure 3:** Gemiddelde positie van advertentie in AH en NRC

# Beeld-tekstverhouding



**Figure 4:** Tekst-beeldverhouding afgezet tegen breedte

## Tekst

---

# Zoeken op woorden in de advertentie

**Delpher**

Kranten

sigaret AND camel

Meer details   Uitgebreid zoeken 

[Soniënie \(20\)](#)

[Verenigde Staten \(35\)](#)

**Soort bericht**

[Advertentie \(497\)](#)

[Artikel \(216\)](#)

[Illustratie met onderschrift \(1\)](#)



**Advertentie**

[sigaret/0 rokers. CAMEL dankt zijn beroemdheid en succes aan zijn geheel eigen smaak. Geniet ook van CAMEL ...](#)

Krantentitel [Algemeen Handelsblad](#)

Datum [27-01-1967](#)

Meer details 

**Krantentitel**

Kies krantentitel...

**Plaats van uitgave**

Kies plaats van uitgave...



**Advertentie**

[sigaret/0 rokers. CAMEL dankt zijn beroemdheid en succes aan zijn geheel eigen smaak. Geniet ook van CAMEL ...](#)

Krantentitel [Algemeen Handelsblad](#)

Datum [09-05-1967](#)

Meer details 

**Herkomst**

Kies herkomst...

**Toevoegingen in Delpher**

[Toon alleen op 13 april 2016 toegevoegde kranten](#)

[Meer informatie](#)



**Advertentie**

[sigaret/0 rokers. CAMEL dankt zijn beroemdheid en succes aan zijn geheel eigen smaak. Geniet ook van CAMEL ...](#)

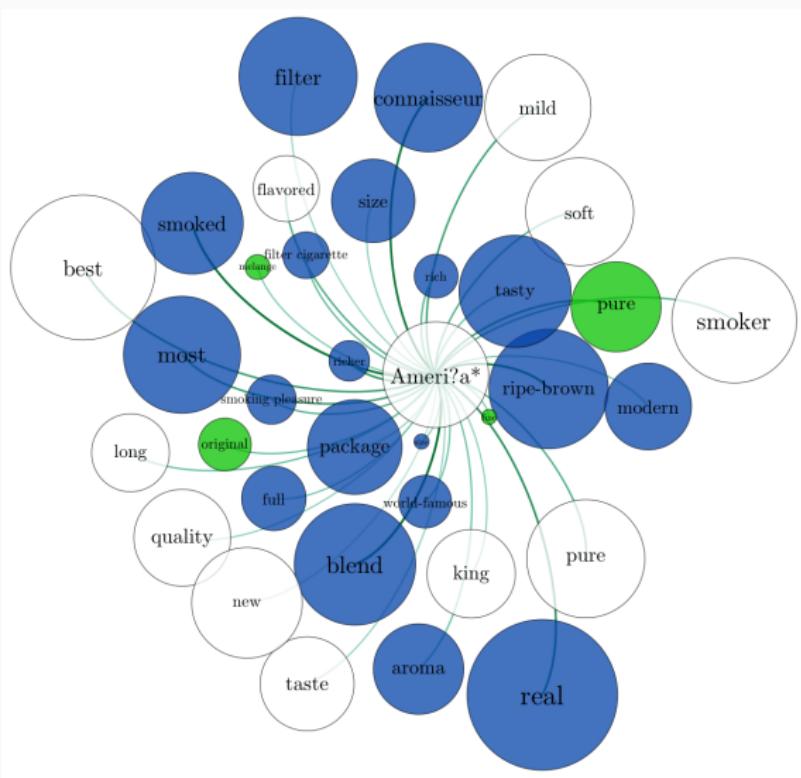
Krantentitel [Algemeen Handelsblad](#)

Datum [26-09-1967](#)

Meer details 

Figure 5: Zoeken in Delpher op 'Sigaret' en 'Camel'

# Tekstanalyse



**Figure 6:** Kenmerken van de Amerikaanse sigaret

# Zoekwoorden genereren

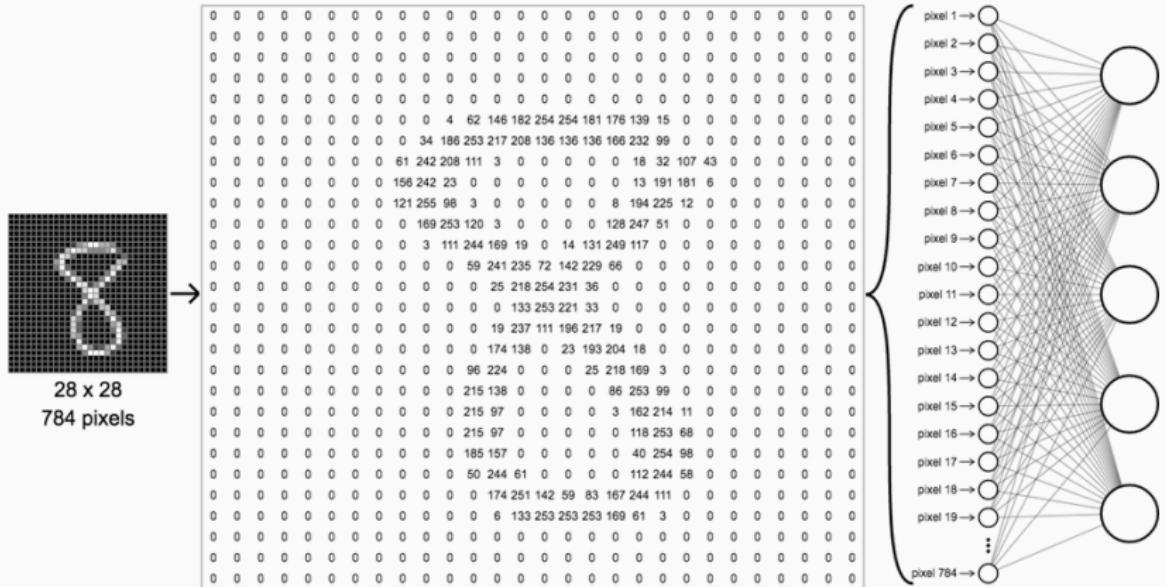


Figure 7: Advertentie C&A

## Beeld

---

## Van plaatje naar data



**Figure 8:** Input layer

# Neuraal network architectuur

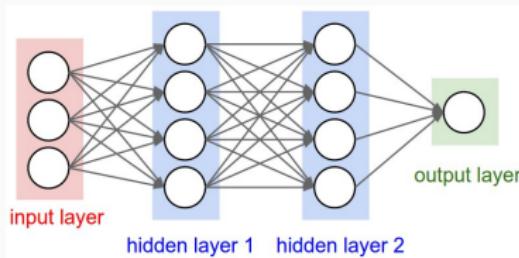


Figure 9: Neuraal netwerk

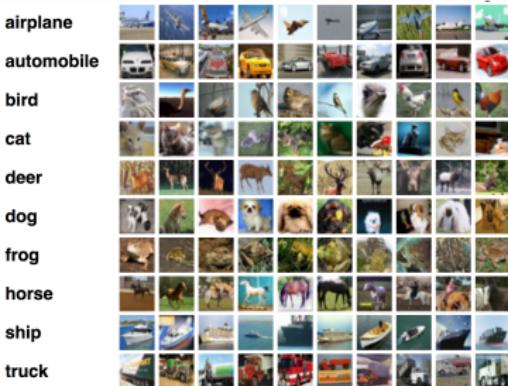


Figure 10: Output layer

# Beeldherkenning



Figure 11: Precisie van herkenning

# SIAMESE

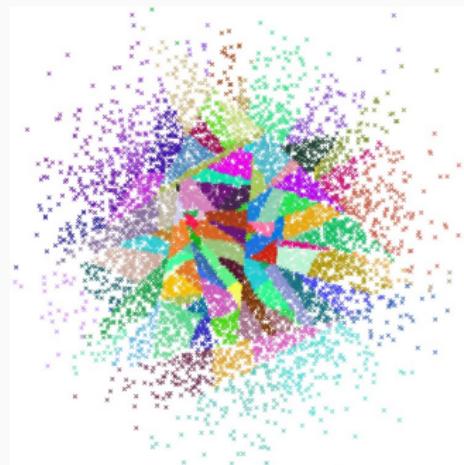
---

## Voorbereiden van de dataset

---

- Dataset: 1.6 miljoen advertenties uit *Algemeen Handelsblad* en *NRC Handelsblad* tussen 1948-1995
- Verwijderen:
  - Kleine afbeeldingen (op basis van metadata)
  - Rubrieksadvertenties
  - Advertenties met voornamelijk tekst
  - Advertenties met auteursrecht
- Resultaat: 426,777 voor de periode 1945-1994

- Afbeeldingen omzetten in vectoren
- Groeperen op basis van 2,048 abstracte visuele aspecten
- Meest gelijkende afbeeldingen vinden (nearest neighbors)
- Tijdslijn met de meest vergelijkbare afbeelding per jaar



**Figure 12:** Annoy:  
Approximate Nearest  
Neighbor

# Tool: SIAMESE

---

- Exploratief door het beeldarchief navigeren op basis van visuele aspecten
- Inzicht krijgen in visuele trends
- Hypotheses genereren, zoekwoorden vinden
- Afbeelding bekijken in context van de krant
- <http://lab.kb.nl/tool/siamese>
- Samen met Juliette Lonij

## Volgende stappen

---

## Wat is er nog meer mogelijk?

- Classifiers trainen voor specifieke producten
- Verrijken collectie
- Grottere set van advertenties inzetten
- Combineren van tekstueel en visueel zoeken

# Vragen/Opmerkingen?

**melvinwevers@icloud.com**

**<http://www.melvinwevers.nl>**

**<http://lab.kb.nl/tool/siamese>**