

# Replikidentifiering av karaktärer från tv-serien The Simpsons med N-Grams

Anton Kindestam [antonki@kth.se](mailto:antonki@kth.se)  
Rasmus Ansin [ransin@kth.se](mailto:ransin@kth.se)

16 december 2015

# Bakgrund

Vi har valt att analysera repliker av The Simpsons-avsnitt för att skapa en klassificerare som kan klassificera vilken yttring som tillhör vilken Simpsons-rollfigur. Vi gör detta genom att skapa en n-gram-databas över yttringar kopplade till rollfigurer. För en given mening genererar vi dess n-gram, och sen väljer vi den rollfigur för vilken denna n-gram-sekvens är sannolikast.

# Parsning av Data

- ▶ Blah1
  - ▶ Blah1.1
- ▶ Blah2

# Vår Modell