

PRESSEMEDDELELSE

# **PRESSEMEDDELELSE - Ny avanceret sprogteknologi læser flydende dansk**

Pressemeddelelse · Rådgivning og teknik



**Infomedia overvåger og analyserer døgnet rundt medieomtaler for virksomheder via 200.000+ mediekilder verden over. Med en ny, avanceret teknologi tilbyder de nu kunderne helt nye muligheder, som samtidig løfter kvaliteten af deres løsninger. Den nye teknologi forstår nemlig det danske sprog bedre, end nogen anden teknologi hidtil har kunnet.**

Lad os sige, at du arbejder i en virksomhed, der gerne vil profilere sig inden for nogle af de 17 verdensmål, FN har formuleret – eksempelvis bekæmpelse af madspild. Du er sikker på,

at jeres løsning virkelig kan gøre en forskel, men der er mange stemmer i debatten omkring madspild, og du vil gerne lægge din energi rigtigt i forhold til at komme igennem med netop jeres budskab.

Et naturligt skridt vil være at bruge en mediantelligence-virksomhed til at holde øje med udviklingen inden for jeres område. Bruger du Infomedias arkiv på over 70 millioner artikler og har daglig medieovervågning af samtlige medietyper verden over, er der ikke noget, du kan misse, tænker du.

Men selv de største spillere inden for medieovervågning og medieanalyser i Danmark har hidtil kunnet misse en vigtig influencer, eller agenda i din søgning. For du har på forhånd skullet vide, hvilke personer og organisationer, du vil holde øje med. Indtil nu. For nu fortæller Infomedias værktøjer dig, hvad du bør holde øje med. Hemmeligheden bag er dansk sprogteknologi og en række avancerede modeller, som er udviklet af Alexandra Instituttet i samarbejde med mediantelligence-virksomheden Infomedia.

*"Den teknologiske udvikling indenfor Natural Language Processing (NLP) har gjort det muligt at tage den teknologiske udvikling inden for det danske sprog til et helt nyt niveau,"* fortæller Joselene Marques, Head of Data Science hos Infomedia.

*"Den nye teknologi gjorde det muligt for os at erstatte en række faste og foruddefinerede regler med en model, der identificerer personer, steder og organisationer, der ikke tidligere blev foreslået i en søgning. I denne tilgang er kvaliteten af de data, der bruges til træning, helt afgørende for succes."*

## **Avanceret sprogteknologi skaber øget relevans**

Teknologien sørger for, at samtaler, eller vigtige aktører, du ikke kender på forhånd, kommer ind på din radar, hurtigere. Kernen i teknologien er baseret på disciplinen 'Named Entity Recognition' (NER), som er en specialiseret gren inden for forskning og udvikling af sprogteknologi.

Med NER kan man træne en computer til at kunne finde eksempelvis personer, lokationer og organisationer/virksomheder i alle slags tekster – dog med den fordel, at computeren også lærer, hvordan de hver især opfører sig grammatisk i en sætning. Det vil sige, at når du fx søger på 'Brøndby' for at finde informationer om, hvilke initiativer de har igangsat i byen i forhold til madspild, så kan du præcisere, at du mener byen 'Brøndby' og dermed eksempelvis undgå at få en række artikler om fodboldklubben Brøndby med i dine resultater.

Simpelthen fordi computeren kan kende forskel på, hvordan der skrives om en lokation og en organisation.

Tidligere kunne man godt søge på ordet '*madspild*' og finde de artikler, der omtaler fænomenet, men derfra var det et tidskrævende, manuelt arbejde i at gå artiklerne igennem for at finde relevante personer og organisationer inden for området. Men de kan nu automatisk blive trukket frem i teksterne, og man kan få en oversigt over, hvem der nævnes oftest, og hvor de opererer fra.

## **Hurtigere og bedre analysearbejde**

Udover at man får et væsentligt forbedret søgeresultat, så ligger der en gevinst i de muligheder, NER åbner op for, når det kommer til at udarbejde analyser til virksomheder på baggrund af deres omtaler i medierne.

Før skulle du vide præcis, hvem du ville holde øje med i mediebilledet. Men nu kan du få det fulde billede af, hvem der dukker op inden for dit område i medierne, og hvem der primært er løbet med taletiden. Måske du har brugt energi på at påvirke politikerne, der sidder på dit felt, mens det i virkeligheden er interesseorganisationerne, der virkelig kunne være en medspiller i forhold til at få dit budskab ud at leve.

Du vil altså kunne få et langt bedre overblik over de emneområder, der påvirker din virksomheds resultater – både hvis du vil snævre det ind og se på enkelte steder eller organisationer eller verdensmål, eller hvis du mere overordnet vil have et billede af, hvem der tager teten på de emner, hvor du selv vil være den førende meningsdanner.

## **Dansk sprogteknologi åbner en vifte af muligheder**

Udover mulighederne med at fremhæve personer, lokationer og organisationer i tekster, kan man træne computeren til at genkende lige det, man har brug for at finde frem til. Det kunne være bestemte produkter eller emner som fx medicin eller symptomer. Kun fantasien sætter grænser.

På samme måde kan anvendelsen af teknologien også bruges til flere ting. Du kan fx både kaste lys over en tekst, eller sørge for at skjule følsomme private informationer. Med mulighederne for persongenkendelse kan du eksempelvis hurtigt finde frem til personfølsomme oplysninger i en tekst, og anonymisere dem. På den måde kan du stadig bruge tekstdataene uden at kompromittere de personer, der nævnes i teksterne.

Alle disse muligheder opstår, når teknologi og tekstdata mødes.

*"Uden store mængder tekstdata kommer man ikke langt med udviklingen af sprogteknologiske løsninger på dansk. Der har samarbejdet med Infomedia været helt ideelt, og har betydet, at vi har kunnet udvikle en hel række af sprogmodeller, som en lang række andre organisationer allerede har haft glæde af via vores open source platform DaNLP. Dansk sprogteknologi er i en rivende udvikling, og vi vil se det brugt i stigende grad, men de færreste virksomheder har ressourcerne, eller adgang til nok data, til selv at udvikle gode modeller. Så samarbejdet her er et skridt i den rigtige retning, når det kommer til at give flere danske virksomheder og organisationer et digitalt løft. Og Infomedia er gået forrest og har anvendt modellerne på inspirerende og effektiv vis", forklarer Anne Bøgh Fangel, Manager – Technologies in Practice, ved Alexandra Instituttet.*

Værktøjerne inden for dansk sprogteknologi ligger frit tilgængelige på Alexandra Instituttets GitHub (<https://github.com/alexandrainst/danlp>). Så arbejder du med dansk sprogteknologi, kan du selv prøve værktøjerne af.

For yderligere information

### **Infomedia A/S**

Niels Bache Larsen, COO

[nba@infomedia.dk](mailto:nba@infomedia.dk)

+45 26 81 10 28

### **Alexandra Instituttet**

Anne Bøgh Fangel, Manager – Technologies in Practice

[anne.b.fangel@alexandra.dk](mailto:anne.b.fangel@alexandra.dk)

+45 22 25 53 22

### **Om Infomedia**

Infomedia-koncernen udvikler og driver digitale abonnementsløsninger til medieovervågning og medieanalyser lokalt og internationalt. Ydelser omfatter bl.a. daglig nyhedsovervågning, social medieovervågning, kriseovervågning, agendaovervågning, interessantovervågning, markedsovervågning, newsroom og journalistdatabaser m.v.

Infomedia har 240 medarbejdere fordelt på kontorer i København, Aarhus, Stockholm, Oslo og Tallinn og ejer desuden teknologiselskabet Opoint Technology, som crawler online nyheder globalt.

Infomedia blev grundlagt i 2002, og ejes i dag 50/50 af Danmarks to største mediehuse, JP/Politikens Hus og Berlingske Media.

## **Om Alexandra Instituttet**

Alexandra Instituttet har siden 1999 været det led, der skubber den nyeste it-forskning helt ud, hvor den skaber værdi i virksomhedernes travle hverdag.

De har hjulpet mere end 1.000 offentlige og private virksomheder med avancerede, effektive, sikre og innovative it-løsninger, der er med til at skabe bæredygtig vækst og velfærd i Danmark.

Deres specialister bruger halvdelen af deres tid på R&D-aktiviteter. Det er deres kunders garanti for, at de har den nødvendige forskningsfaglige dybde, som kræves for at rådgive om og udvikle fremtidssikre, digitale teknologier og services.

*Kilde / Logo / Grafik: Alexandra Instituttet*

Online: <https://pressemeddelelse.dk/pressemeddelelser/pressemeddelelse-ny-avanceret-sprogteknologi-laeser-flydende-dansk/>

Pressemeddelelsen er videreformidlet af Pressemeddelelse.dk. Pressemeddelelse.dk er ikke ansvarlig for indholdet. Spørgsmål til indholdet rettes til Alexandra Instituttet.