



Aalborg Universitet

AALBORG UNIVERSITY  
DENMARK

## Datafantasi - Tre case virksomheder - 2023

Ørngreen, Rikke; Munk, Anders Kristian; Knudsen, Sara Paasch; Søltoft, Johan Irving; Hansen, Helene Husted; Jacomy, Mathieu; Knudsen, Asger Gehrt

*Creative Commons License*  
Ikke-specificeret

*Publication date:*  
2023

*Document Version*  
Anden version

[Link to publication from Aalborg University](#)

*Citation for published version (APA):*

Ørngreen, R., Munk, A. K., Knudsen, S. P., Søltoft, J. I., Hansen, H. H., Jacomy, M., & Knudsen, A. G. (2023, feb. 13). Datafantasi - Tre case virksomheder - 2023.

### General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal -

### Take down policy

If you believe that this document breaches copyright please contact us at [vbn@aub.aau.dk](mailto:vbn@aub.aau.dk) providing details, and we will remove access to the work immediately and investigate your claim.

# Data fantasi

## – Tre case virksomheder

Fra projektet "Styrkelse af SMV'ernes datafantasi"  
af

Rikke Ørngreen, Anders Kristian Munk, Sara Paasch Knudsen, Johan Irving  
Søltoft, Helene Husted Hansen, Mathieu Jacomy, Asger Gehrt Olesen

Aalborg Universitet 2023

De tre virksomheders  
arbejde med  
datafantasi?



## Erhverv Norddanmark

Erhverv Norddanmark er en netværksorganisation med mere end 650 medlemmer inden for handel, industri, service og liberale erhverv. En del af deres arbejde består i at lave arrangementer for deres medlemmer.

Kan vi lære  
noget nyt?

## HUSK AT TJEKKE VORES ERHVERVSKALENDER

- Du finder den under "arrangementer"  
på [www.erhvervnorddanmark.dk](http://www.erhvervnorddanmark.dk)



Hvert år afholder Erhverv Norddanmark en række forskellige arrangementer for deres medlemmer.

De har oversigter over, hvem der deltager i arrangementerne, hvem medlemmerne er (type, antal ansatte mv.) og beskrivelser af, hvad arrangementerne handler om. Derudover findes der ekstern data på deres medlemmers hjemmesider. De vil gerne se, om det er muligt at trække mere viden ud af både intern og ekstern data.

# Eksperiment #1

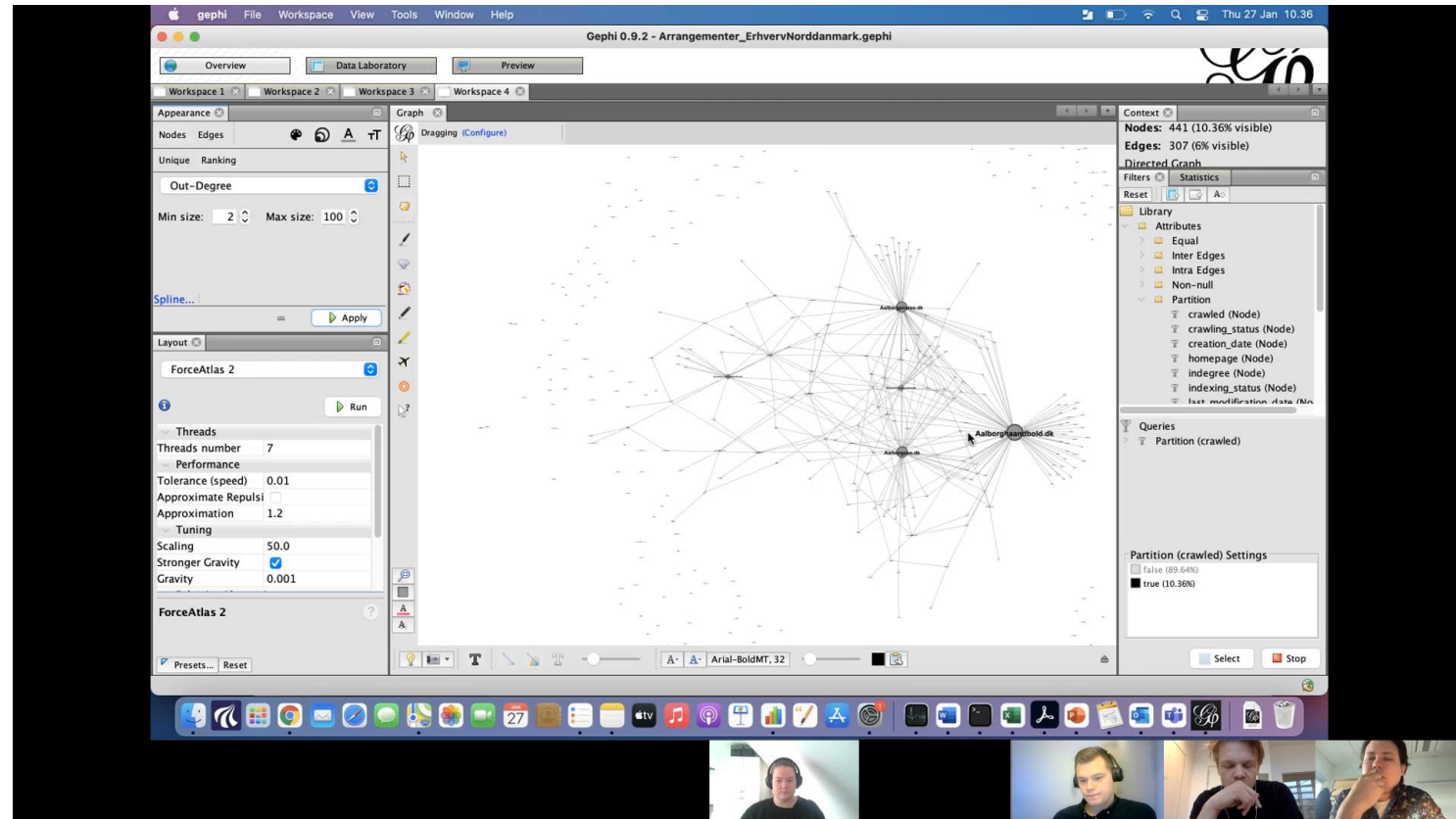
*Kan vi blive klogere på vores medlemmers relationer til hinanden ved at undersøge deres hjemmesider?*

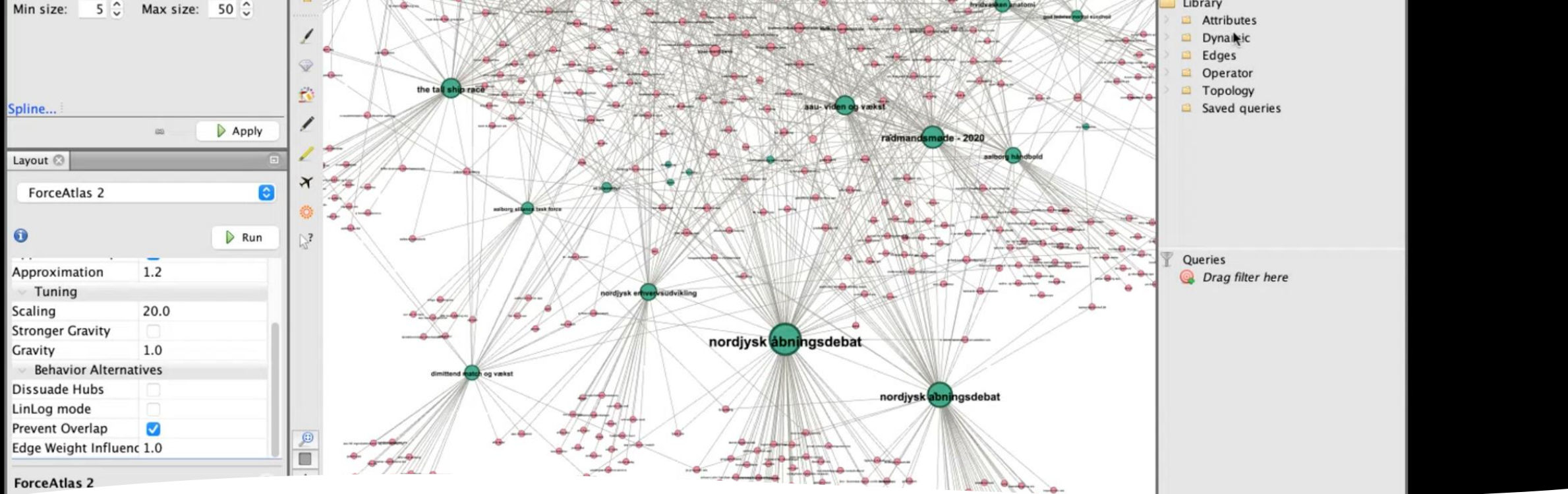
Illustrationen på videoen viser hvordan medlemmernes hjemmesider linker til hinanden.

Det ses, at nogle medlemmer henviser mere til hinanden end andre, og at der er en håndfuld lokale aktører, som medlemmerne generelt henviser meget til.

I videoen diskuteres, hvad det fortæller os, og om den viden overhovedet er relevant for Erhverv Norddanmarks arbejde.

Erhverv Norddanmark giver efterfølgende udtryk for, at de er mere interesserede i at hjælpe deres medlemmer med at skabe relationer ved at møde hinanden ansigt til ansigt. På den baggrund blev det besluttet i stedet at prioritere data om, hvordan medlemmerne deltager i Erhverv Norddanmarks arrangementer.





## Eksperiment #2

*Hvem deltager i vores arrangementer?*

*Del 1 - I den første runde arbejdes med en netværksvisualisering.*

Illustrationen viser et link mellem Erhverv Norddanmarks arrangementer og personer som har deltaget i disse arrangementer i perioden 2007-2022.

Netværksvisualiseringen opleves abstrakt og kompleks. Der udtrykkes behov for mere klassiske og genkendelige visualiseringer som grafer og tidslinjer.

# Eksperiment #2

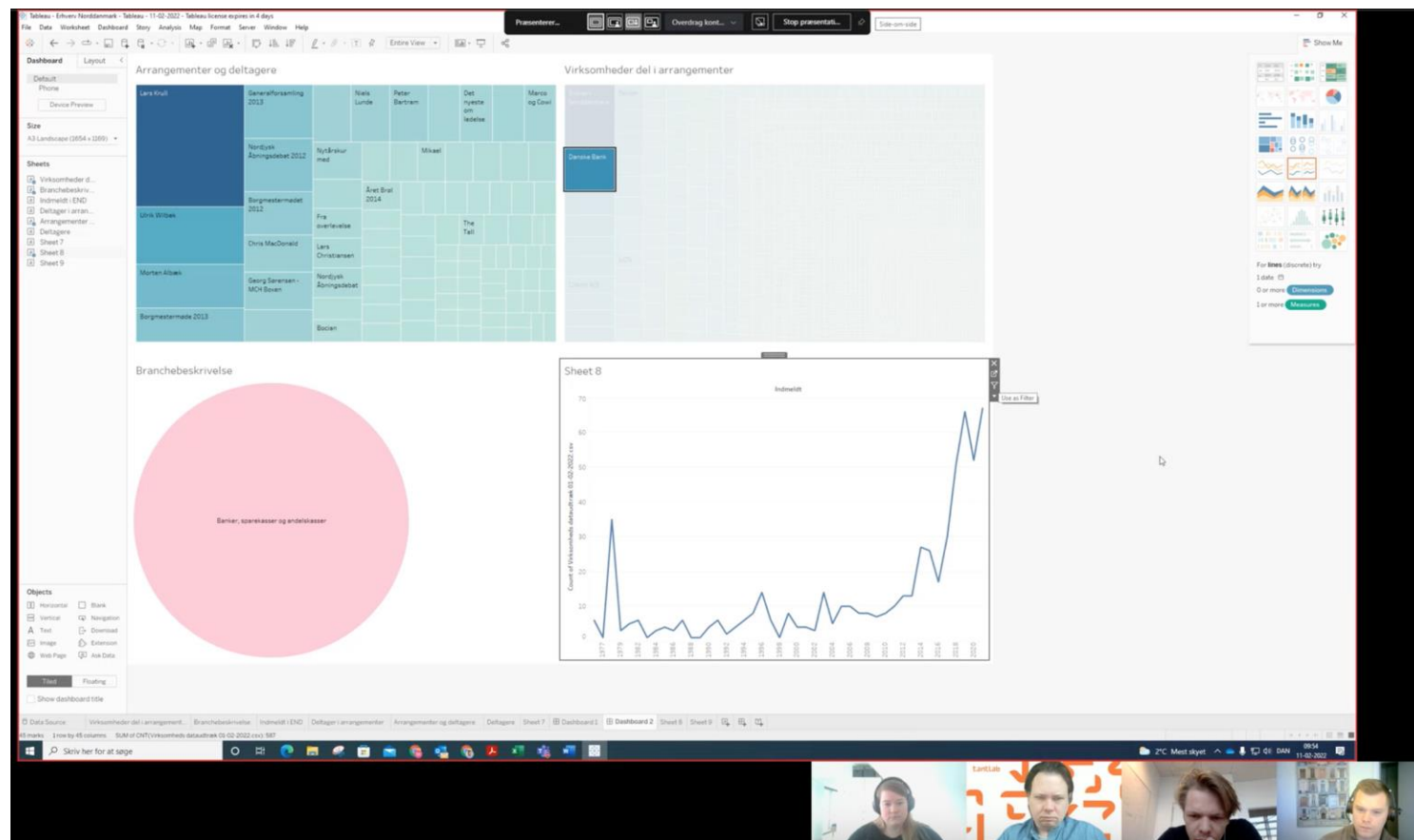
## *Hvem deltager i vores arrangementer?*

*Del 2 – I den anden runde arbejdes med et dataoverblik.*

Illustrationen viser et interaktivt dashboard, der giver overblik over arrangementerne over tid, antal deltagere, deltagerens organisationer og deres geografiske placering.

Dashbordet giver eksempelvis mulighed for at udforske hvilke typer af arrangementer, der har haft mange deltagere, eller hvor deltagerne kommer fra.

Disse typer af visualiseringer var mere genkendelig for deltagerne, der oplevede det som mere brugbart og dermed værdifuldt.









# Programmer benyttet:

## Eksperiment 1:

**Hyphe** blev brugt til at scrape hyperlinks på alle Erhverv Norddanmarks medlemsorganisationers hjemmesider. Derefter blev **Gephi** benyttet til at visualiserer netværket

## Eksperiment 2:

*Del 1*

**Python** blev brugt til at rengøre excel-ark fra 2007-2022. Derefter blev **Table2net**, brugt til at lave bipartite netværk, som blev visualiseret i **Gephi**.

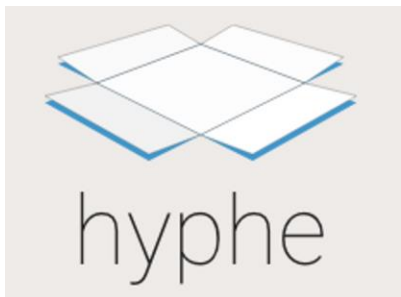
## Eksperiment 2:

*Del 2*

Det rengjorte excel-ark blev også visualiseret ved brug af **Tableau**, hvor der blev lavet interaktive dashboards, der forbandt arrangements, år, deltagere, organisationer og geografisk placering.

## Eksperiment 3:

**CorText**, blev benyttet til at lave natural language processing (NLP) på alle arrangements beskrivelserne fra årrækken 2007-2022.



<https://hyphe.medialab.sciences-po.fr/>



<https://gephi.org/>



<https://www.python.org/>



Table 2 Net

<https://medialab.github.io/table2net/>



<https://www.tableau.com/>



<https://managerv2.cortext.net/>

# Hvad er datafantasiproblemet og løsningen hos Erhverv Norddanmark

---

- Erhverv Norddanmark opsamler og gemmer data om medlemmerne, deres organisationer, arrangementer og deltagerne i arrangementerne, men der er ikke en ensartet praksis for, hvordan det gøres.
- Målet er ikke at forandre deres nuværende praksis, men at optimere det, de i forvejen er gode til. Det betyder blandt andet, at man ikke umiddelbart tænker på de kvalitative arrangementsbeskrivelser eller medlemmernes hjemmesider som data. **Analyse af arrangementsdata har primært handlet om det, der kan tælles.**
- **Datafantasien udvides i første omgang da det bliver klart, hvordan en ændret strukturering af de tælbare arrangementsdata gør det muligt automatisk at skabe opdaterede visuelle overblik** over emner, der før har krævet manuel håndtering. Her har hverken spørgsmål eller visuelle formater ændret sig fra det, Erhverv Norddanmark allerede gør.
- **I anden omgang udvides datafantasien på en anden måde. Det sker, når Erhverv Norddanmark begynder at overveje, hvordan en automatisk tekstanalyse af deres arrangementsbeskrivelser kunne spille en rolle i onboarding af nye medarbejdere** (i dette tilfælde en ny chef). Her opdager man nye spørgsmål, man kunne stille til data.
- Virksomheden er blevet udfordret på måden at tænke data på og har taget de første skridt i forhold til at se på data på en ny måde.

# Bobach

Bobach er en stålvirksomhed, der blandt andet producerer stålaltaner og -gelændere.

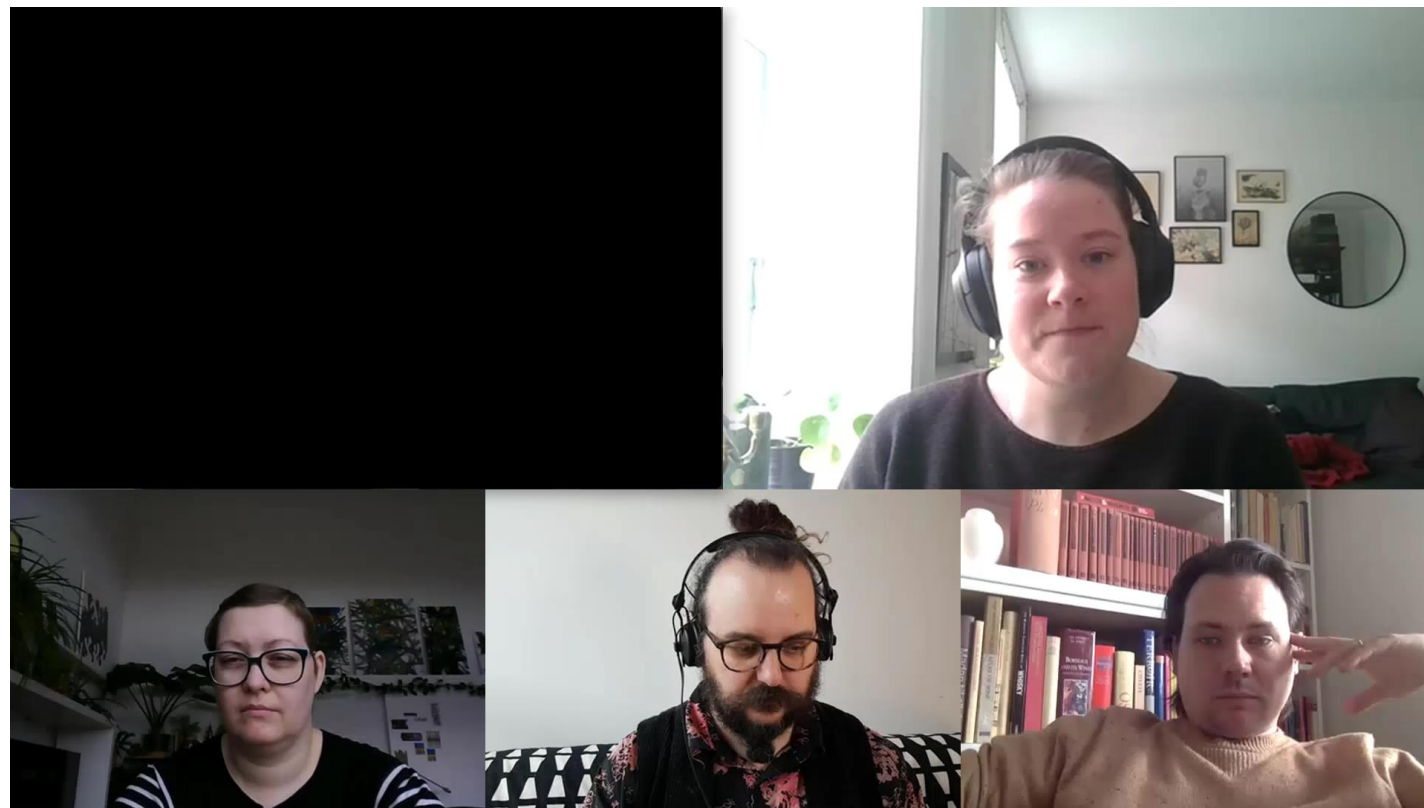
En del af arbejdet består i at vise mulige kunder billeder af, hvordan disse produkter kan se ud i færdige projekter.



# Kan en maskine gøre mit arbejde?

Medarbejderens udfordring er, at hun er alene om at bestyre det store billedarkiv.

Andre i organisationen kan ikke finde rundt i arkivet, derfor spørger de altid hende, når de skal finde et billede.



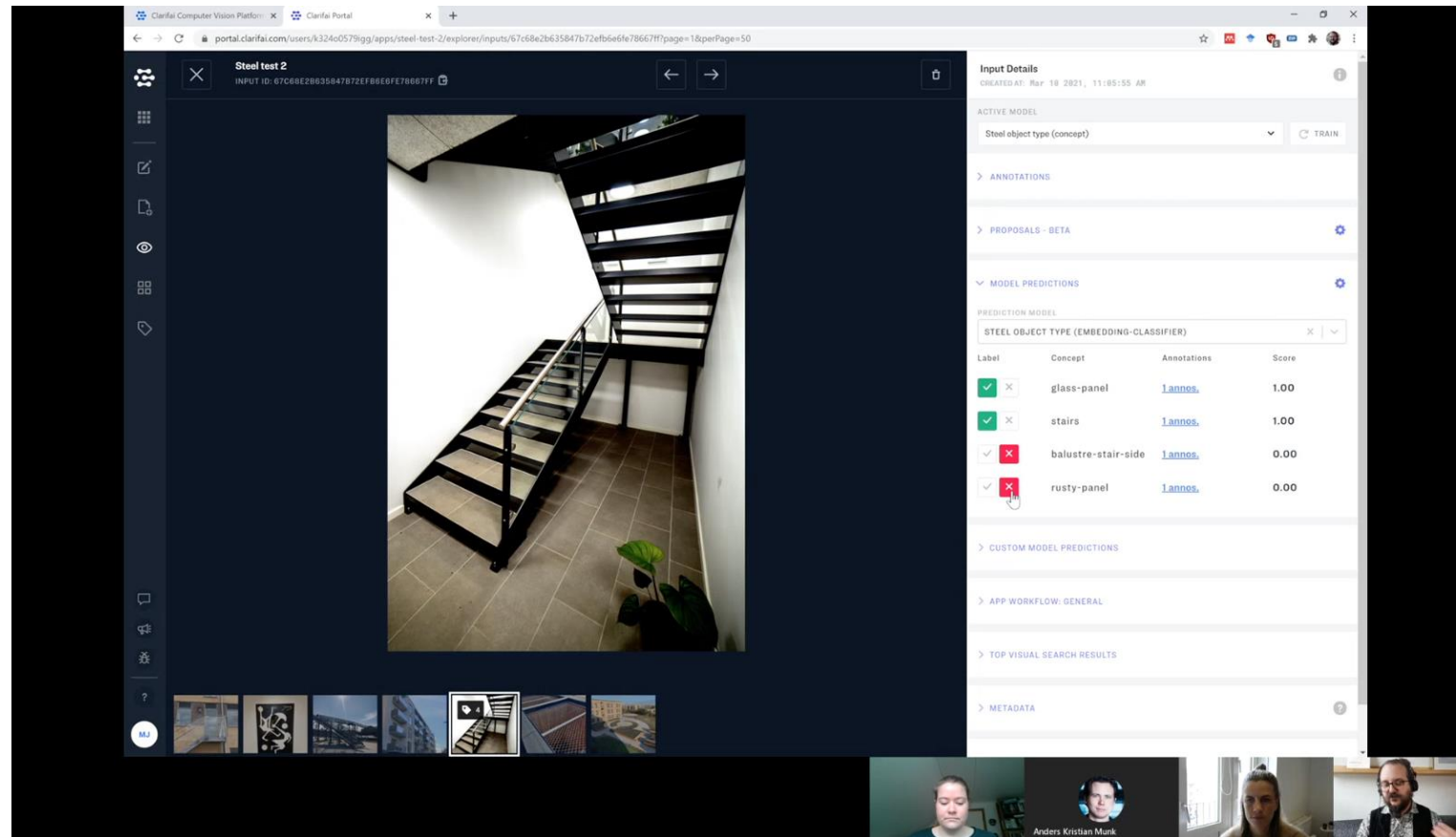
# Eksperiment #1

## *Kan en maskine gøre mit arbejde?*

En superviseret maskinlæringsmodel (Clarifai) bliver trænet på baggrund af den måde medarbejderen allerede kategoriserer sine egne billeder på.

Modellen forsøger at efterligne medarbejderens kategoriseringer og viser sig på den baggrund at kunne skelne altaner fra gelændere m.m.

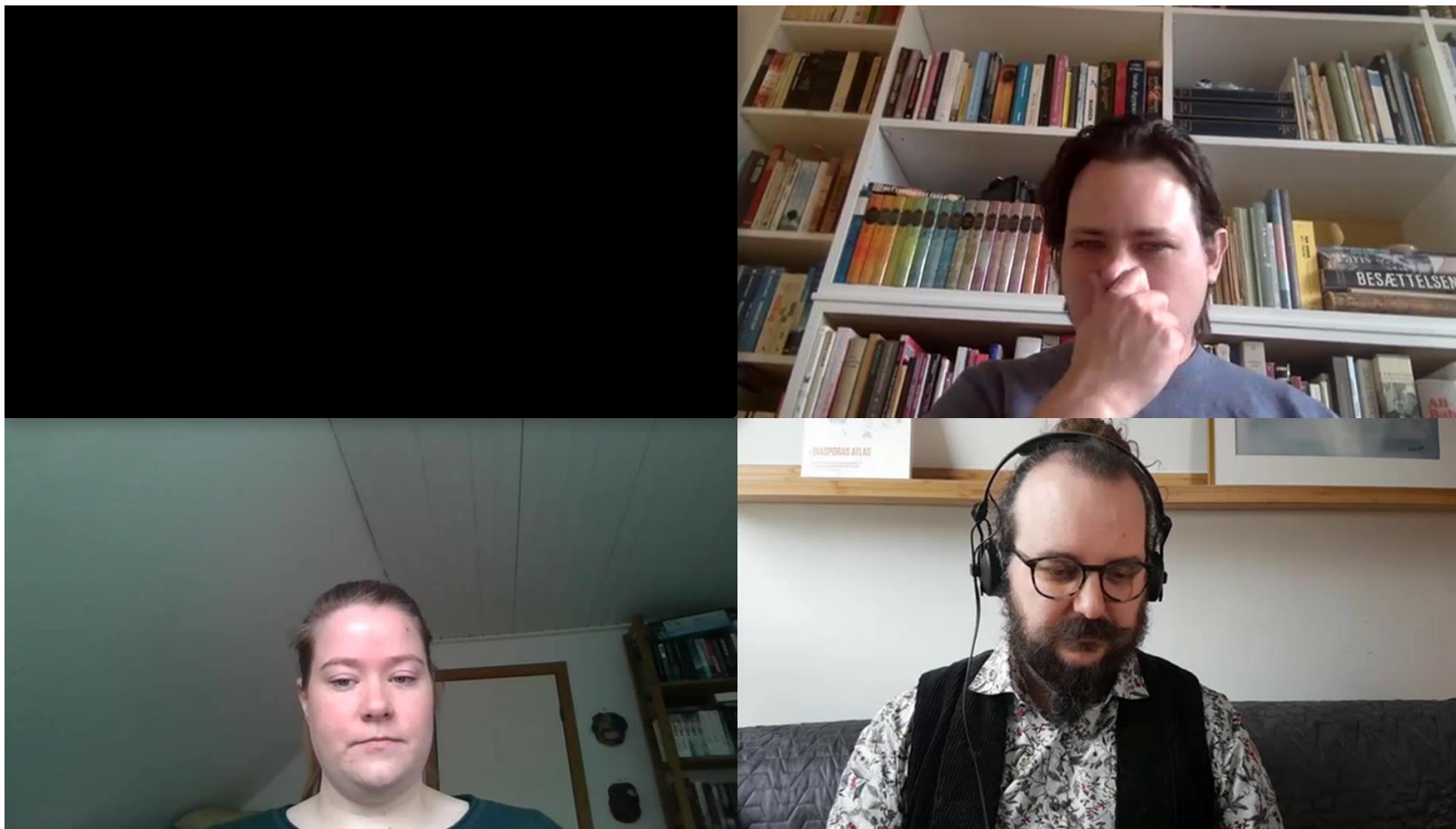
Resultatet giver dog også anledning til at diskutere om modellen kan genskabe den detaljegråd, som medarbejderen selv arbejder med, og som hun oplever som nødvendig.



The screenshot displays the Clarifai Computer Vision Platform interface. The main window shows a photo of a staircase with a glass railing. On the right, the 'Input Details' panel is visible, showing the active model 'Steel object type (concept)'. Below this, the 'MODEL PREDICTIONS' section displays a table of results:

Label	Concept	Annotations	Score
<input checked="" type="checkbox"/>	glass-panel	<a href="#">1 anno.</a>	1.00
<input checked="" type="checkbox"/>	stairs	<a href="#">1 anno.</a>	1.00
<input checked="" type="checkbox"/>	balustre-stair-side	<a href="#">1 anno.</a>	0.00
<input checked="" type="checkbox"/>	rusty-panel	<a href="#">1 anno.</a>	0.00

At the bottom of the interface, a video call window shows four participants, including one identified as 'Anders Kristian Munk'.



Muligheder og  
begrænsninger

Maskinen har ikke samme  
detaljegråd og forståelse.



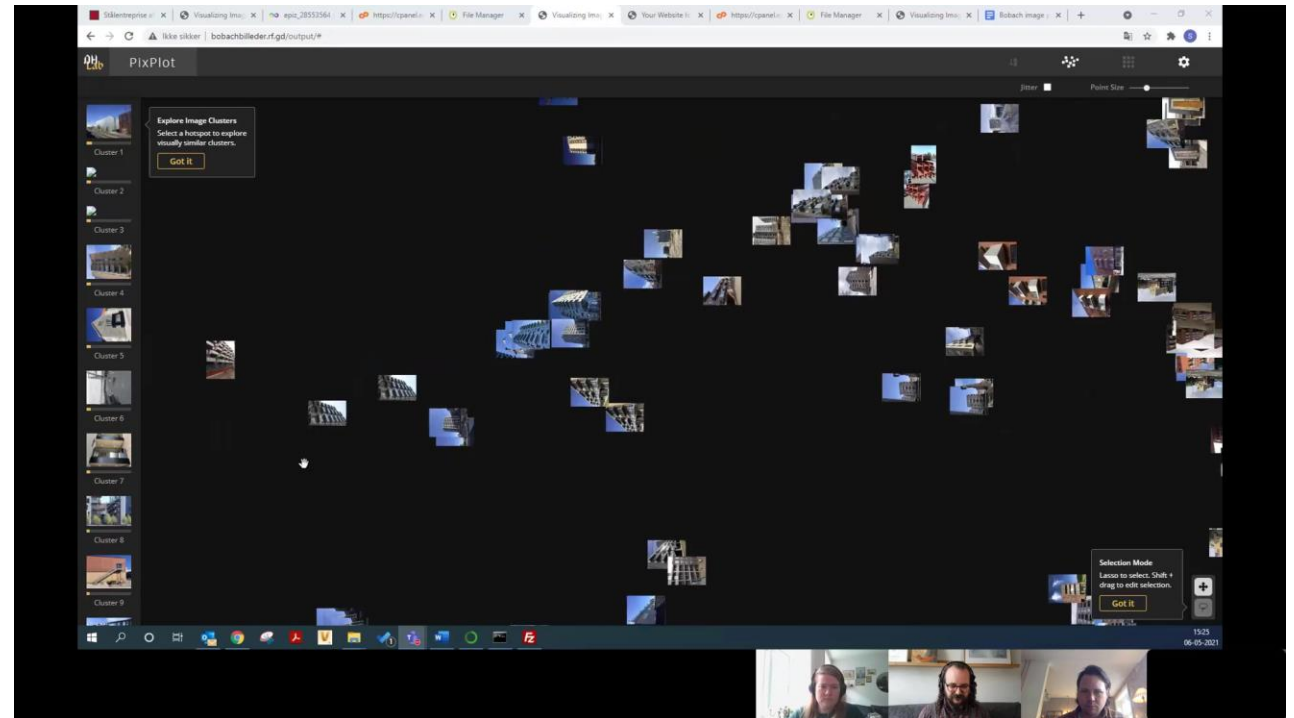
## Eksperiment #2

### *Supervisereret maskinlæring kan ikke opnå nødvendig detaljegrad*

Strategien bliver ændret. I stedet for at forsøge at efterligne detaljegraden i medarbejderens kategorisering af billederne lader vi en usuperviseret model klassificere billederne på egen hånd.

I stedet for at bede maskinen om at fortælle os, hvad der er på billederne, beder vi den om at finde ligheder imellem billederne med henblik på at vise mønstre.

Spørgsmålet er nu, om det kan fungere som en alternativ navigationsvej ind i billedarkivet.



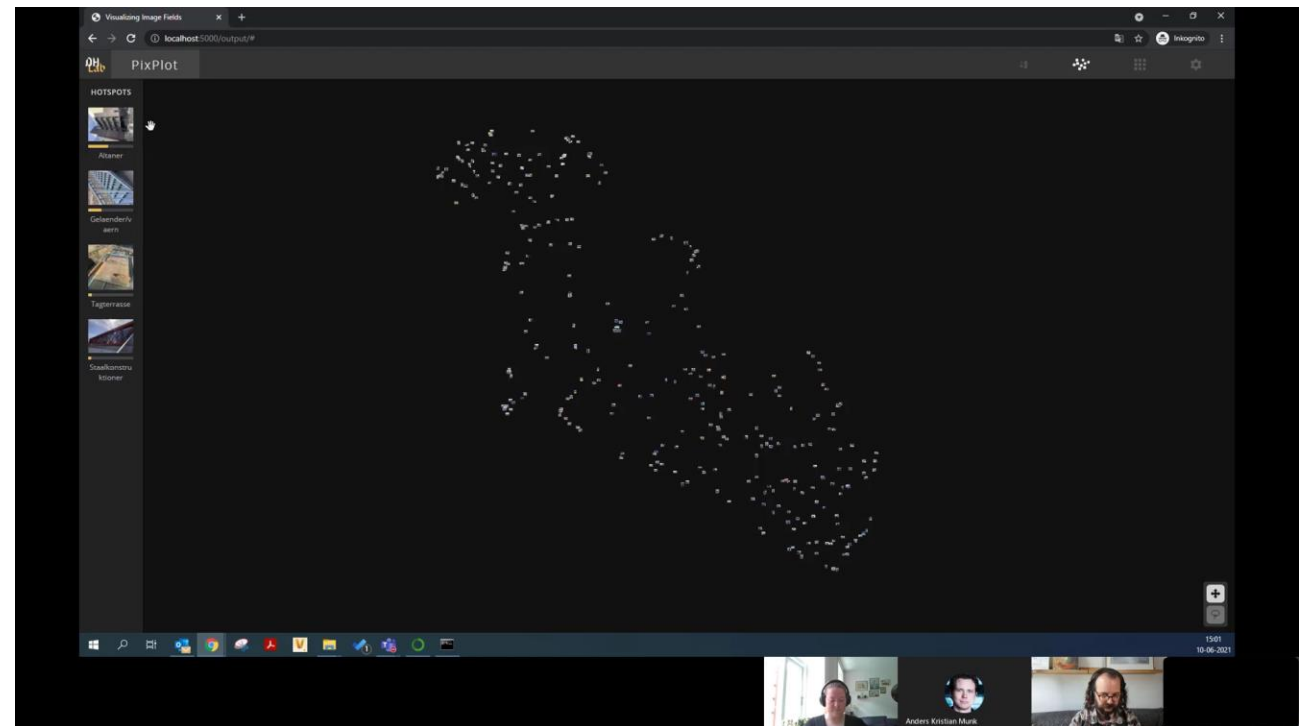
På videoen ses den første dialog om, hvordan medarbejderen ville begynde at navigere i de mønstre som den usuperviserede model gav.

## Eksperiment #2

### *Nyt overblik og ny navigation gennem billedarkivet*

Under eksperimentet opstod der en diskussion om, hvorvidt det er maskinens opgave at ramme præcist det, man ønsker, eller om der er behov for en anden tilgang.

Det gav anledning til en anden måde at danne overblik på, som kan bruges til at navigere i billedarkivet.



På videoen er medarbejderen i gang med at vise rundt i det visuelle overblik maskinen har produceret. Det er i denne løsning muligt for Bobachs medarbejdere selv at finde billederne. Det vil være op til brugeren at vælge de billeder, man har brug for, men maskinens bidrag er at skabe den orden, der gør det muligt at overskue alle billederne visuelt og samtidigt.

The update: add 4 images to ~350 images and re-train the PixPlot

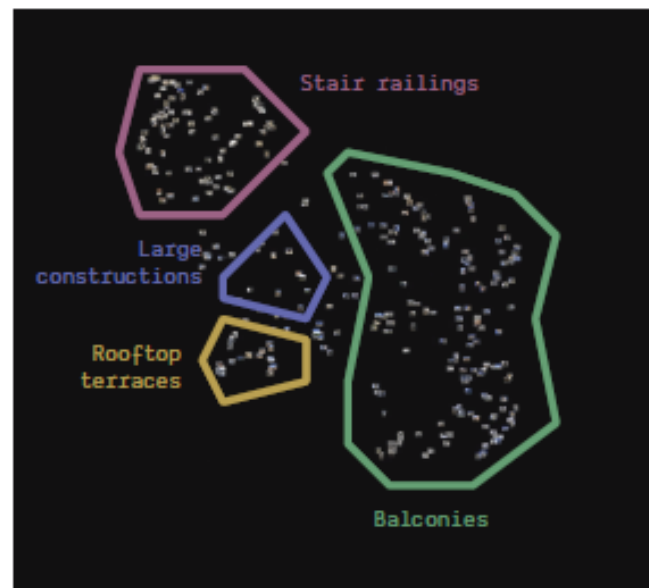
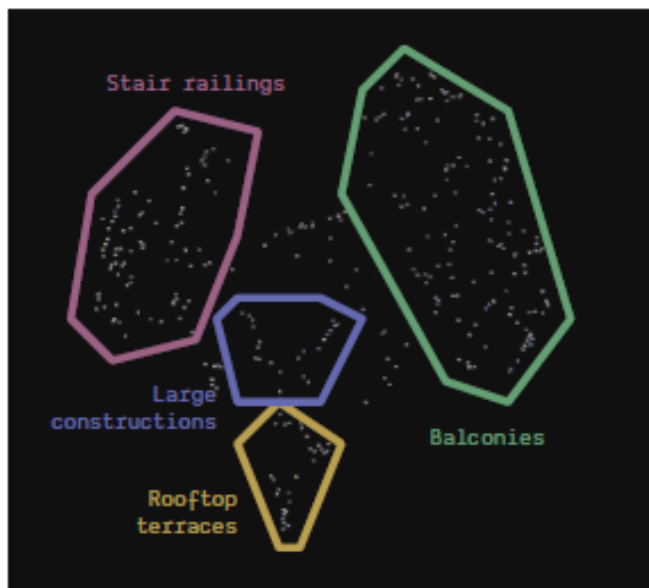
Before update



After update



Clusters



## Kortlægning af billeder fra projekter baseret på deres visuelle ligheder og forskelle

Kortlægningen anvender maskinlæring til at analysere billederne og grupperer dem således, at billeder, der ligner hinanden, placeres i nærheden af hinanden i de samme klynger. I en efterfølgende analyse har vi annoteret hver enkelt klynge i et forsøg på at beskrive de visuelle fællestræk, der knytter billederne sammen.



# Programmer benyttet:

## Eksperiment 1:

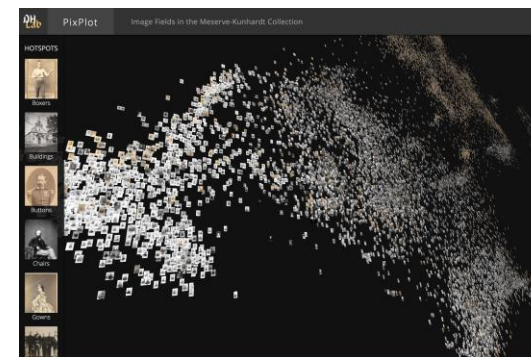
En superviseret maskinlæringsmodel – **Clarifai** blev brugt til at identificerer motiver på billederne, som medarbejderen allerede navigerer efter.



<https://www.clarifai.com/>

## Eksperiment 2:

En usuperviseret maskinlæringsmodel – **Pixplot** blev brugt til at gruppere billederne efter visuelle ligheder.



<http://pixplot.com/>

# Hvad er datafantasiproblemet og løsningen hos Bobach

---

- Deltagerne fra Bobach har ikke svært ved at forestille sig, hvad problemet er. **Den eksisterende procedure for at finde relevante billeder skaber en uholdbar flaskehals.**
- Den første intuition er at få en maskine til at efterligne den eksisterende procedure med henblik på at automatiser den.
- Det viser sig, at maskinen ikke kan blive ligeså god som den etablerede løsning, fordi algoritmen ikke kan trænes til at opnå samme detaljegråd i billedgenkendelse som medarbejderen.
- **Udvidelse af datafantasiens består i at droppe tanken om at efterligne den eksisterende metode og i stedet omorganisere arbejdet på en måde, der udnytter det, maskinen faktisk kan finde ud af.**
- Det kræver en accept af ændrede arbejdsgange, hvor det vil være kollegaerne der skal søge billeder frem, til gengæld fjernes flaskehals problematikken.

# Bolighuset Werenberg

Werenberg er et møbelhus, der både har en fysisk forretning, en hjemmeside og en Instagramkonto, hvor de sælger og markedsfører 30 forskellige brands.

En del af arbejdet handler om at vise deres varer frem på de bedst mulige måder med udgangspunkt i, hvad der appellerer til deres kunder.



# Kan vi finde nye måder at bruge Instagram i vores markedsføring?

Werenberg bruger i stigende grad influencere til at markedsføre deres produkter på Instagram. De bruger en blanding af Instagrams metrikker for engagement (hvor mange nås per opslag), deres erfaringer og deres personlige vurdering af billedernes stil og æstetiske udtryk til at vælge, hvilke influencere de skal arbejde sammen med.

Werenbergs udfordring handler om at få bedre overblik over, *hvem* det er interessant at nå ud til, og ikke bare *hvor mange* de når ud til.



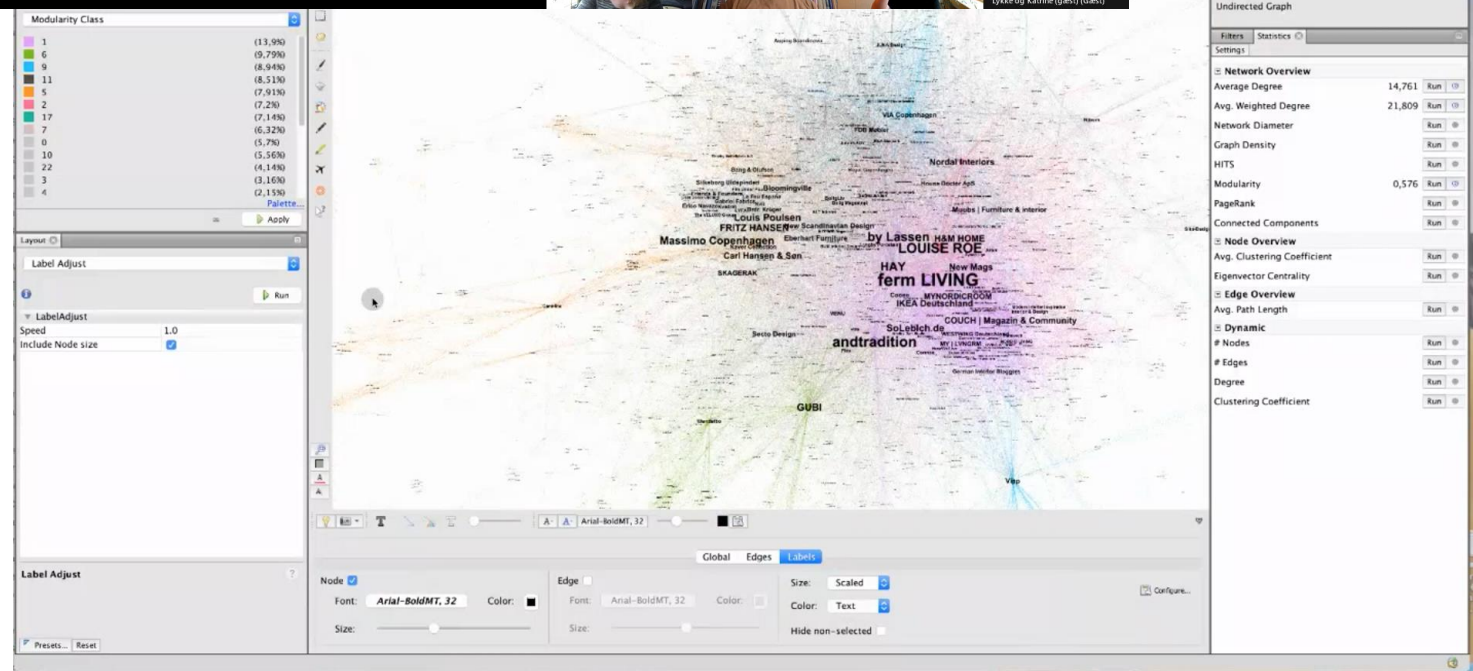
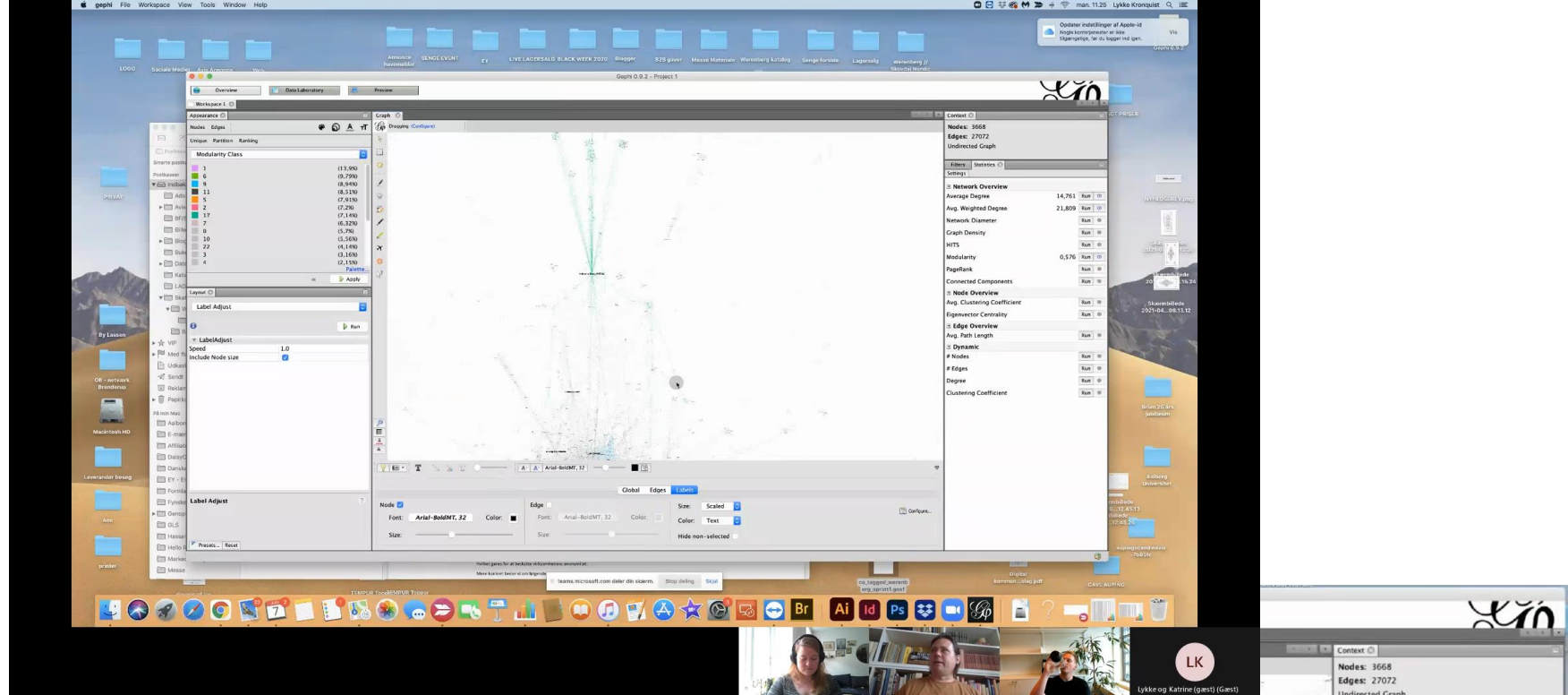
# Eksperiment #1

## Hvordan bliver vores brands vist frem på Instagram?

I videoen undersøger vi resultatet af vores første eksperiment med at kortlægge, hvordan brugere tagger de 30 brands.

Illustrationen viser hvilke brands, der bliver tagget i de samme opslag.

Det viser sig, at der er store forskelle på hvordan brands bliver tagget og hvem der tagger dem. Det fører til en overvejelse om, hvem Werenberg når ud til, når de tagger forskellige brands i deres opslag. Det handler ikke bare om, hvor mange de får fat i, men hvilke sammenhænge, de bliver set i.





# Eksperiment #2

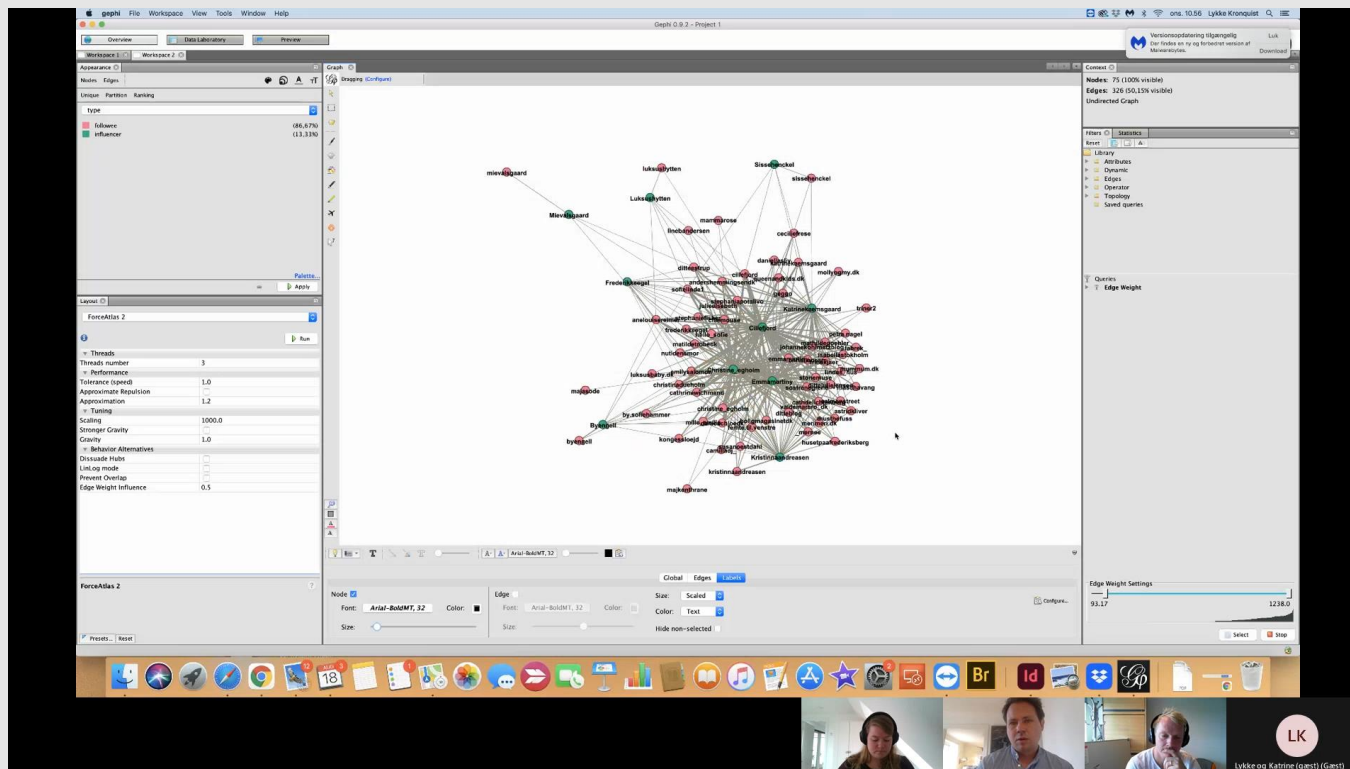
## Hvilke følgere har vores brands og influencers til fælles?

Den første kortlægning rejste spørgsmålet om, hvorvidt og hvordan de influencers, som Werenberg bruger til markedsføring, følges af de samme brugere som de brands, Werenberg sælger.

I videoen diskuteres resultatet af en ny kortlægning, der inkluderer en række influencers, som Werenberg har udvalgt.

Illustrationen viser hvilke brands og influencers, der deler følgere med hinanden.

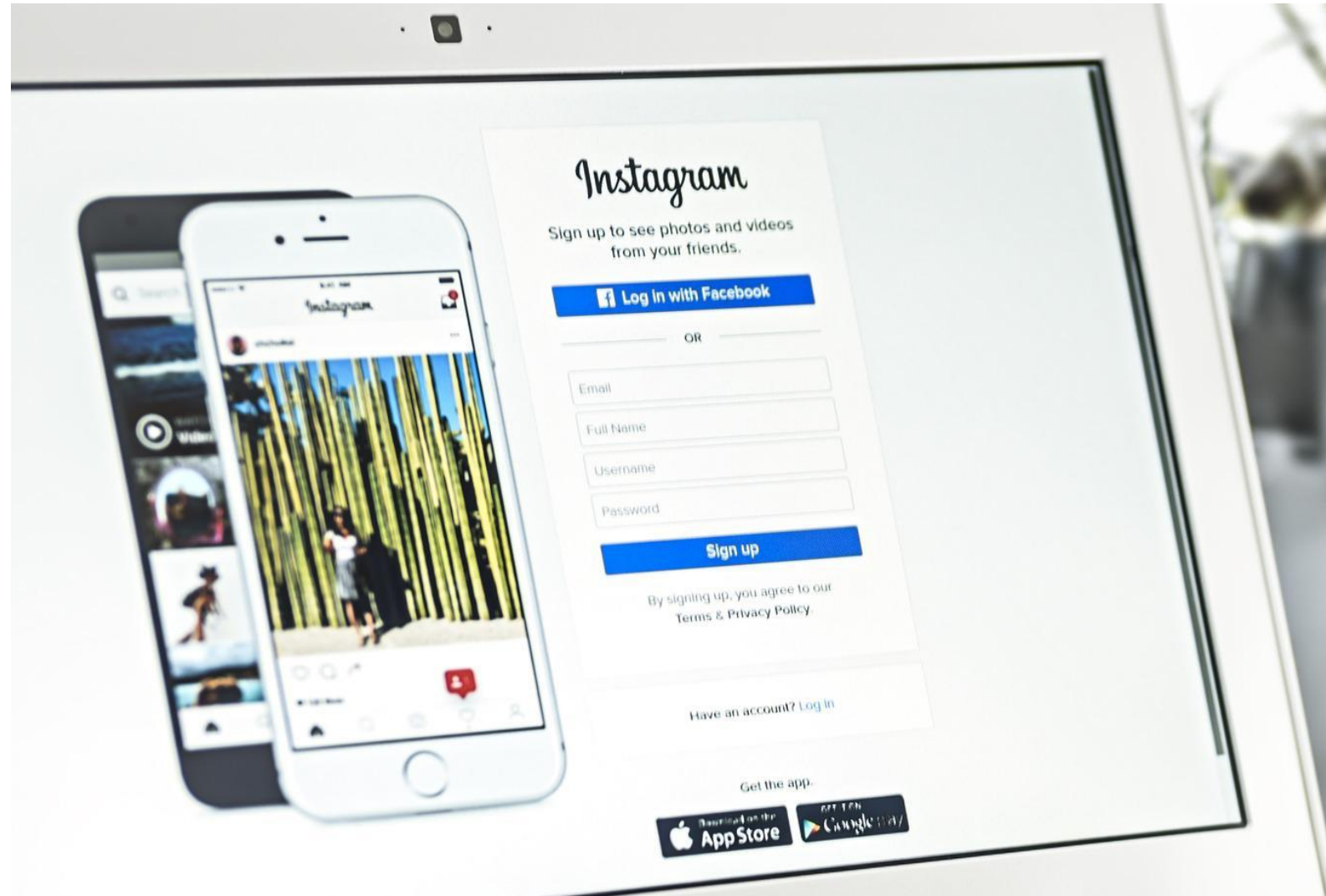
Det bliver klart, at nogle influencers har mange af de samme følgere som bestemte af Werenbergs brands. Andre influencers ikke har så mange af de samme følgere eller har følgere tilfælles med andre af Werenbergs brands.



# Eksperiment #3

*Kan vi følge effekten af en kampagne på Instagram?*

Eksperiment #1 og #2 gav anledning til at diskutere, om man kunne følge effekten af en konkret kampagne med tre influencere på Instagram på en ny måde. I den sidste kortlægning sammenlignedes, hvordan brugerne på Instagram interagerede med indhold fra henholdsvis Werenberg og de tre influencere i perioden op til og umiddelbart efter kampagnen.

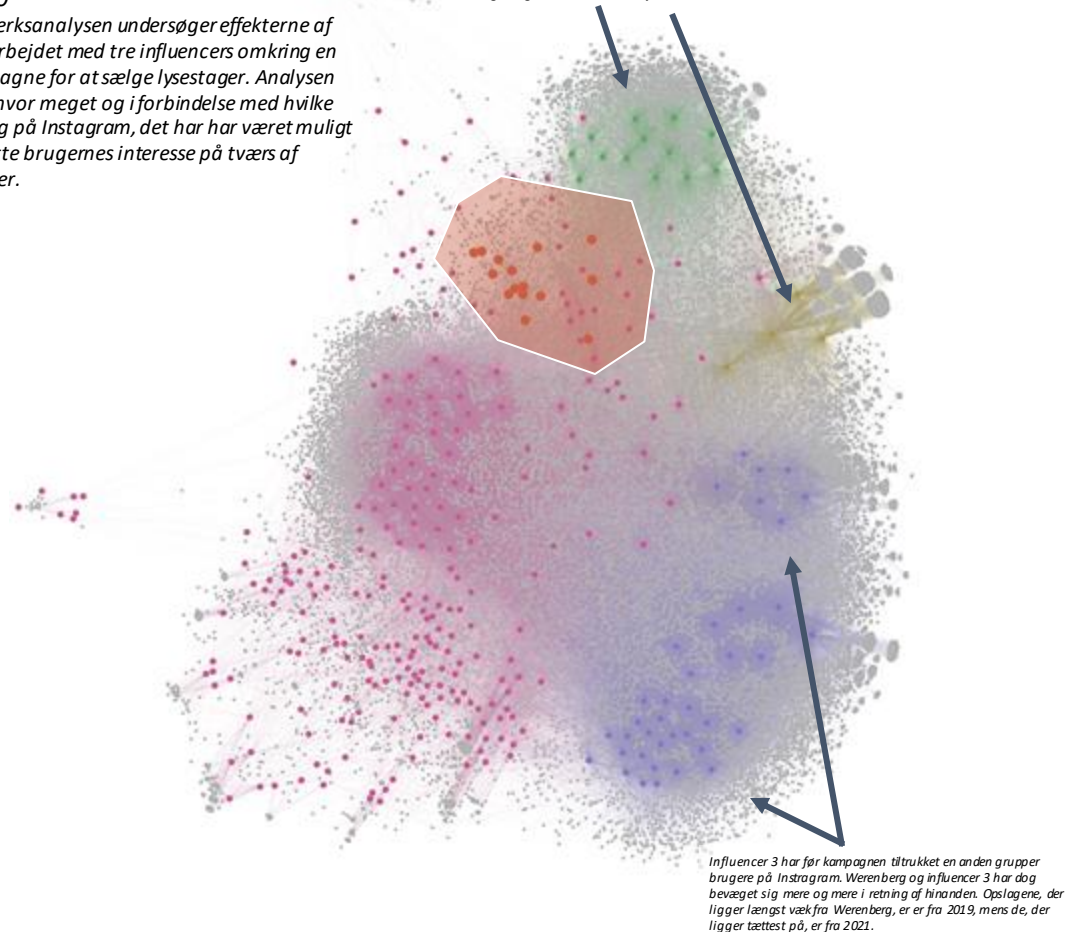


## Før kampagnen

1 august, 2017 - 30 august, 2021, 12:00

Netværksanalysen undersøger effekterne af samarbejdet med tre influencers omkring en kampagne for at sælge lystager. Analysen viser hvor meget og i forbindelse med hvilke opslag på Instagram, det har været muligt at flytte brugernes interesse på tværs af profiler.

Werenberg har størst brugeroverlap med influencer 1 og 2. Dette skyldes formentlig at Werenberg tidligere har haft samarbejder med disse to influencere.



Influencer 3 har før kampagnen tiltrukket en anden gruppe brugere på Instagram. Werenberg og influencer 3 har dog bevæget sig mere og mere i retning af hinanden. Opslagene, der ligger længst væk fra Werenberg, er fra 2019, mens de, der ligger tættest på, er fra 2021.

## Efter kampagnen

30 august, 2021, 12:00 - 10 oktober, 2021

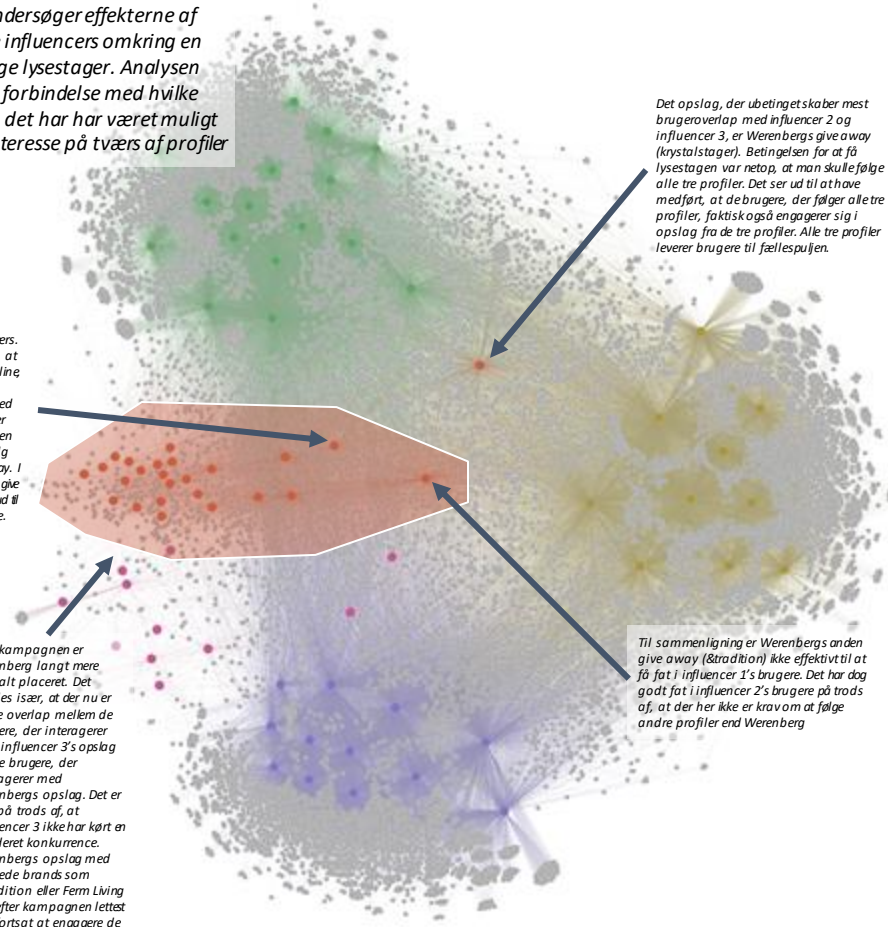
Netværksanalysen undersøger effekterne af samarbejdet med tre influencers omkring en kampagne for at sælge lystager. Analysen viser hvor meget og i forbindelse med hvilke opslag på Instagram, det har været muligt at flytte brugernes interesse på tværs af profiler.

Det er ikke kun selve konkurrencen der skaber fælles engagement på tværs. Opslaget, der annoncerer, at lystagerne er til salg online, har nogenlunde samme effekt. Det harmonerer med statistik fra Instagram, der viser større trafik til profilen på opslaget om give away, i forhold til eksponering er give away opslaget dog nået ud til dobbelt så mange brugere.

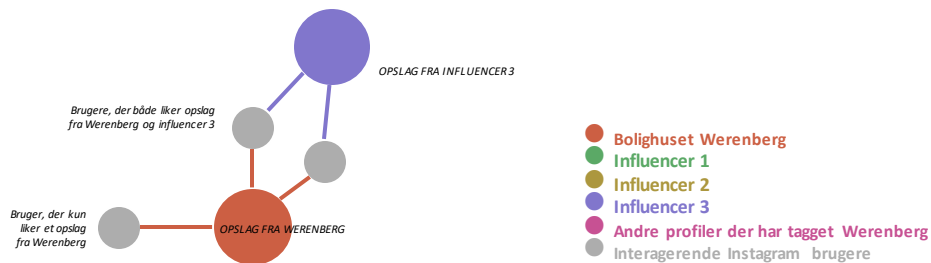
Efter kampagnen er Werenberg langt mere centralt placeret. Det skyldes især, at der nu er større overlap mellem de brugere, der interagerer med influencer 3's opslag og de brugere, der interagerer med Werenbergs opslag. Det er sket på trods af, at influencer 3 ikke har kørt en decideret konkurrence. Werenbergs opslag med taggede brands som &tradition eller Fem Living har efter kampagnen lettest ved fortsat at engagere de nye brugere.

Det opslag, der ubetinget skaber mest brugeroverlap med influencer 2 og influencer 3, er Werenbergs give away (krystalstager). Betingelsen for at få lystager var netop, at man skulle følge alle tre profiler. Det ser ud til at have medført, at de brugere, der følger alle tre profiler, faktisk også engagerer sig i opslag fra de tre profiler. Alle tre profiler leverer brugere til fællespuljen.

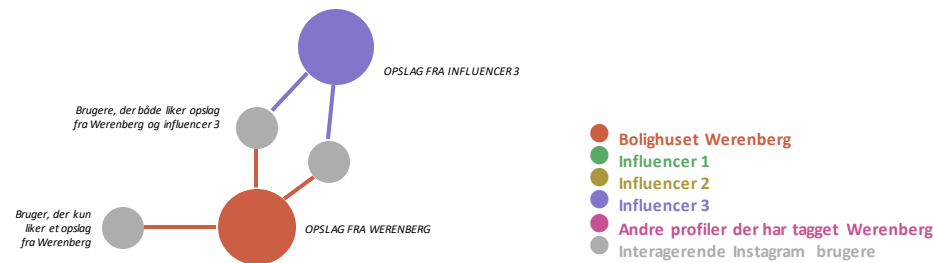
Til sammenligning er Werenbergs anden give away (&tradition) ikke effektivt til at få fat i influencer 1's brugere. Det har dog godt fat i influencer 2's brugere på trods af, at der her ikke er krav om at følge andre profiler end Werenberg.



### Hvordan er netværket lavet?

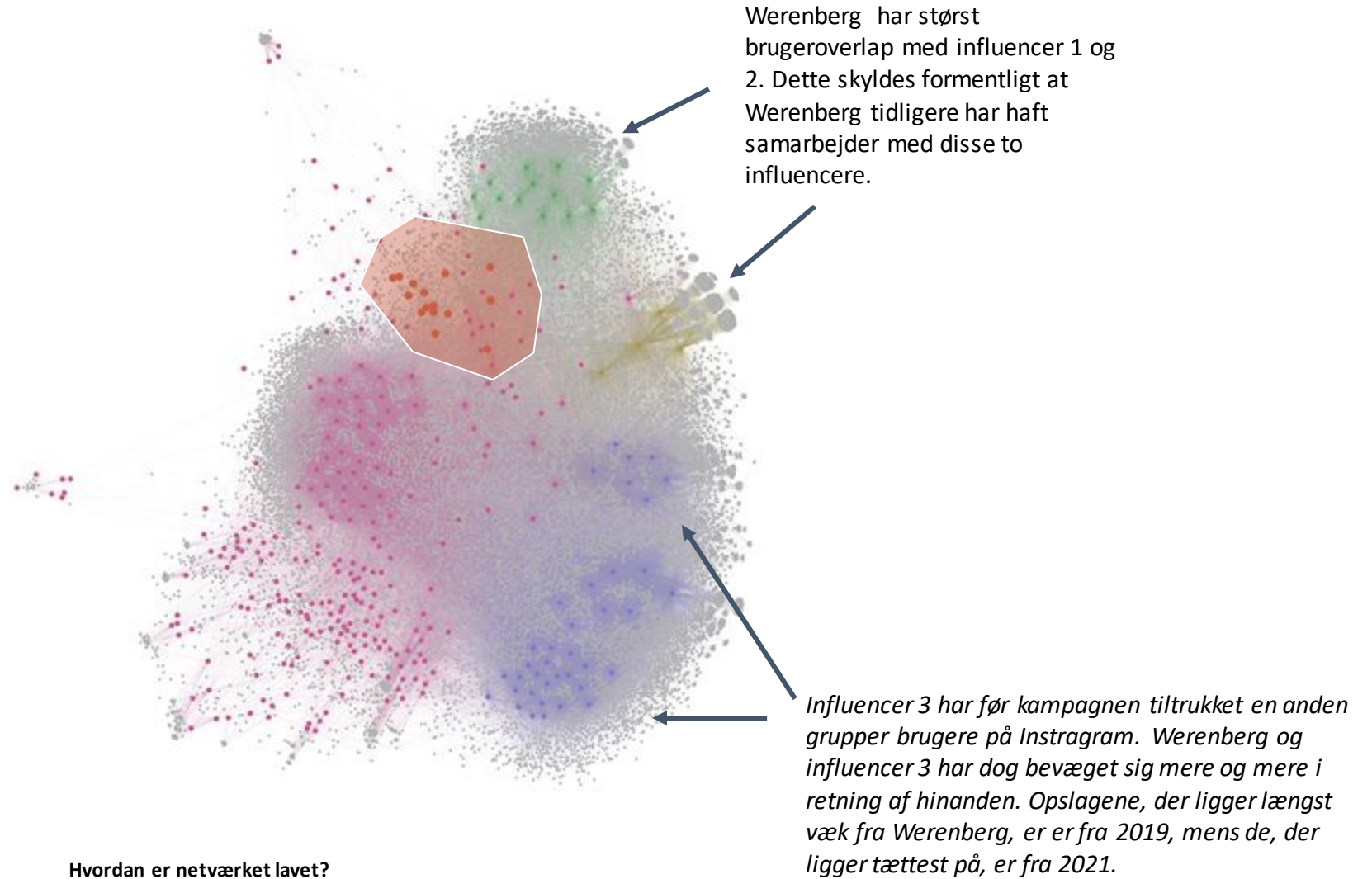


### Hvordan er netværket lavet?

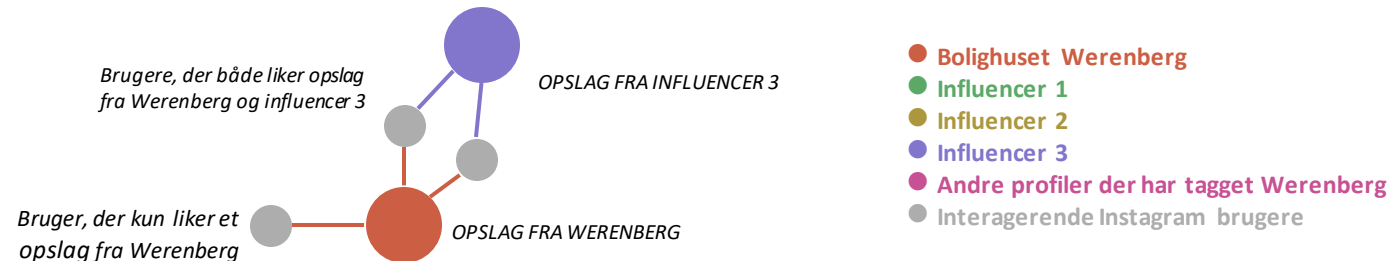


# Før kampagnen 1 august, 2017 - 30 august, 2021, 12:00

Netværksanalysen undersøger effekterne af samarbejdet mellem Werenberg, og tre influencers omkring en kampagne for at sælge lysestager. Analysen viser hvor meget og i forbindelse med hvilke opslag på Instagram, det har været muligt at flytte brugernes interesse på tværs af profiler.



## Hvordan er netværket lavet?



# Efter kampagnen

30 august, 2021, 12:00 -  
10 oktober, 2021

Netværksanalysen undersøger effekterne af samarbejdet mellem Werenberg, og tre influencers omkring en kampagne for at sælge lysestager. Analysen viser hvor meget og i forbindelse med hvilke opslag på Instagram, det har været muligt at flytte brugernes interesse på tværs af profiler.

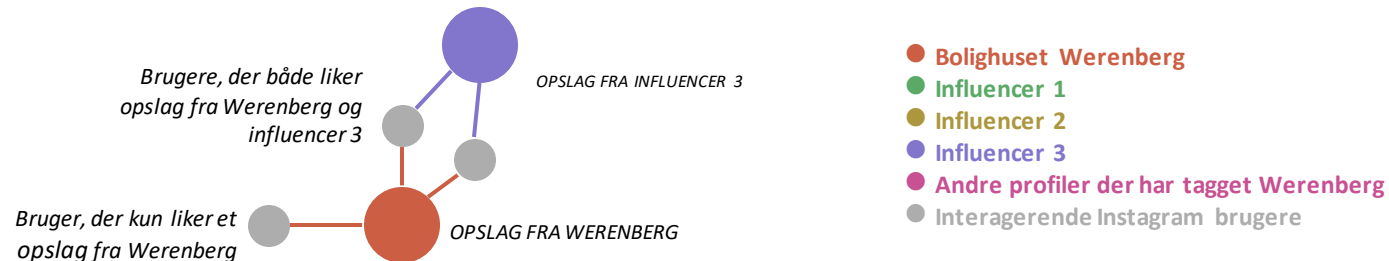
Det er ikke kun selve konkurrencen der skaber fælles engagement på tværs. Opslaget, der annoncerer, at lysestagerne er til salg online, har nogenlunde samme effekt. Det harmonerer med statistik fra Instagram, der viser større trafik til profilen på opslaget om online salg end opslaget om give away. I forhold til eksponering er give away opslaget dog nået ud til dobbelt så mange brugere.

Det opslag, der ubetinget skaber mest brugeroverlap mellem influencer 1 og influencer 2, er Werenbergs give away (krystalstager). Betingelsen for at få lysestagen var netop, at man skulle følge alle tre profiler. Det ser ud til at have medført, at de brugere, der følger alle tre profiler, faktisk også engagerer sig i opslag fra de tre profiler. Alle tre profiler leverer brugere til fællespuljen.

Efter kampagnen er Werenberg langt mere centralt placeret. Det skyldes især, at der nu er større overlap mellem de brugere, der interagerer med influencer 3's opslag og de brugere, der interagerer med Werenbergs opslag. Det er sket på trods af, at influencer 3 ikke har kørt en decideret konkurrence. Werenbergs opslag med taggede brands som &tradition eller Ferm Living har efter kampagnen lettest ved fortsat at engagere de nye brugere.

Til sammenligning er Werenbergs anden give away (&tradition) ikke effektivt til at få fat i influencer 1's brugere. Det har dog godt fat i influencer 2's brugere på trods af, at der her ikke er krav om at følge andre profiler end Werenberg

## Hvordan er netværket lavet?



Hvordan kan indsigterne fra projektet påvirke Werenbergs arbejde med Instagram?

– den videre færd med influencermarketing

Deltagelse i projektet har givet en udvidet forståelse af Instagram som markedsplads.

Deltagerne fra Werenberg er blevet klar, at det er vigtigt for dem at vurdere influencere på andet end stil og antal følgere. Det giver anledning til et nyt fokus på, hvilke typer følgere, man når ud til, og hvilke sammenhænge, man bliver set i.



# Programmer benyttet:

## Eksperiment 1:

Et **Python** script blev brugt til at scrape data fra Werenbergs Instagramprofil. Den hentede data visualiseres gennem redskabet **Gephi** til at skabe forskellige netværks.



<https://www.python.org/>

## Eksperiment 2:

Et nyt **Python** script blev brugt til at scrape data fra Werenbergs profil fra Instagram, som udover de 30 brands også indeholder information om 11 udvalgte influencers. **Gephi** bruges til at skabe et netværk over de 30 brands, og hvordan influencerne og deres følgere placerer sig i forhold til disse brands.

## Eksperiment 3:

Et **Python** script blev brugt til at scrape Instagram lige før kampagnens start og lige efter kampagnens afslutning. **Gephi** blev brugt til at skabe to netværk, hvor man kan se før og efter kampagnen.



<https://gephi.org/>

# Hvad er datafantasiproblemet og løsningen hos Bolighuset Werenberg?

---

- Bolighuset Werenberg har en etableret praksis omkring Instagram og influencere, der bygger på faglig intuition og erfaringer fra branchen. De har også kvantitativt praksis, hvor de måler engagement ved at kigge tal fra Instagram, der viser hvor mange brugere de når ud til.
- Instagram tilbyder et interface, der understøtter den praksis. De leverer både engagementsmetriker og mulighed for at browse opslag.
- Udvidelse af datafantasien består i at indhente og visualisere data fra Instagram på nye måder og på den måde også understøtte en anden måde at forstå, hvem det er man taler til.
- **De gik fra at tælle engagerede brugere til at skelne mellem interessefællesskaber. Dermed går man også fra at vurdere effekten af en kampagne som et spørgsmål om, hvor mange man får fat i, til også at være et spørgsmål om, hvem man får fat i.**
- Det kræver tekniske kompetencer, der kan være svære at oparbejde. Det kan løses ved at kompetenceudvikle eksisterende medarbejdere, ansætte nye eller outsource opgaven.
- Det har dog også en resulteret i en øget bevidsthed om, hvilke typer spørgsmål man kan forvente at få svar på.



# Generelle betragtninger

- **Der er en række velkendte dataudfordringer med velkendte løsninger, som ikke kræver udvidelse af SMV'ernes datafantasi.** Det kan for eksempel handle om at automatisere meget strukturerede manuelle opgaver, eller om at bruge etablerede redskaber og metoder til at svare på etablerede spørgsmål.
- Omvendt peger casestudierne på, at en udvidelse af **datafantasien giver mulighed for at udnytte data på nye måder og danne grundlag for at optimere eksisterende praksis eller skabe nye løsninger.**

# Generelle betragtninger

- Et vigtigt element i **at arbejde med datafantasi er at være nysgerrig, stille spørgsmål og arbejde med processer, hvor man ikke ved på forhånd hvilke resultater det vil give.** Casene har vist, at det er muligt at arbejde konstruktivt med datafantasi, også når de forskellige SMV'er har haft forskellige udgangspunkter og vilkår for organisationen, for medarbejderne og deres data.
- **Casene har vist at datafantasi kræver refleksion, forberedelse, tid og aktivt data arbejde. Det kan dog være svært for en SMV at skabe rum til dette,** i en hverdagskontor-logik med daglig drift. Det er muligt at lave mindre projekter, der udvikler dele af datafantasi, mens fuld udnyttelse af mulighederne både kræver udvikling af medarbejderen evne og kompetencer, fokus på data kvalitet, adgang og fleksibilitet, samt organisationens tilgang til og prioritering af tilgangen.