

Issue Observatory Search - Hurtig Guide

En trin-for-trin guide til at lave en komplet søgning, scraping og netværksanalyse.

Kan tilgås på <http://212.27.13.34:3111>

!Din browser vil muligvis advare dig mod at gå videre til websted. Dette skal du ignorere og bare gå videre. Nogle gange skal man klikke på “Advanced” for at kunne vælge at gå videre.

Brugernavn og password skal du have fået. Ellers skriv til jakobbk@ruc.dk

Trinvejs Proces

1. Login

Log ind med dit brugernavn og password.

2. Opret Søgning

1. Klik på “**Search Jobs**” → “**New Search Session**”
2. Udfyld:
 - **Session Name:** “Dansk Fodbold Sponsorering 2024”
 - **Search Engine:** Serper
 - **Search Terms** (ét søgeord pr. linje):
danske fodboldsponsorer
fodboldklubber sponsorater
virksomheder fodbold danmark
 - **Results per query:** 10
3. Klik “**Create Search Session**”
4. Vent til søgningen er færdig (viser “completed”)

Resultat: 40 URLs fundet fra 25 forskellige domæner

3. Lav Søge-Netværk

1. Klik på “**Networks**” → “**Generate Network**”
2. Udfyld:
 - **Network Name:** “Søgeords Netværk”
 - **Network Type:** “Search Results Network”
 - **Search Session:** Vælg din søgning
 - **Include Search Terms:**
 - **Include Domains:**
3. Klik “**Generate Network**”
4. Vent til den er færdig (10-30 sekunder)

5. Klik **“Download GEXF”** og gem filen
6. !!! Din browser vil muligvis forsøge at blokere download. Se hvad den gør og giv eksPLICIT tilladelse.

Resultat: 29 noder (4 søgeord + 25 domæner), 40 forbindelser

4. Vurder Søgeresultater

1. Gå tilbage til din søgning (klik på navnet)
2. Klik **“View Results”** for hver søgning
3. Tjek:
 - Er URLs relevante?
 - God blanding af kilder?
 - Autoritative hjemmesider?

Vurdering: Fandt virksomhedshjemmesider, sportsorganisationer, nyhedsartikler - god nok kvalitet til at fortsætte.

5. Start Scraping

1. På din søgningsside, klik **“Start Scraping”**
2. Udfyld:
 - **Job Name:** “Fodbold Sponsorering Scrape”
 - **Scraping Depth:** “Level 1 - Search results only”
 - **Excluded Domains** (valgfrit):
linkedin.com
facebook.com
3. Klik **“Start Job”**
4. Følg fremgangen (opdateres automatisk hvert 3. sekund)

Forventet tid: 10-12 minutter for 40 URLs **Resultat:** 37/40 URLs scraped succesfuldt (32 fuld succes + 5 brugte snippet som fallback)

6. Lav Indhold-Netværk

1. Når scraping er **færdig** (eller delvist færdig hvis cancelled), gå til **“Networks”** → **“Generate Network”**
2. Udfyld:
 - **Network Name:** “Sponsorering Begreber”
 - **Network Type:** “Scraping Job Network”
 - **Scraping Job:** Vælg dit scraping job
 - **Include Nouns:**
 - **Include Domains:**

- **Minimum Word Frequency:** 3
3. Klik **“Generate Network”**
 4. Vent (3-4 minutter)
 5. Klik **“Download GEXF”** og gem filen
 6. !!! Din browser vil muligvis forsøge at blokere download. Se hvad den gør og giv eksplicit tilladelse.

Tip: Du kan lave netværket selvom scraping blev cancelled - det bruger bare de sider der blev scraped succesfuldt.

Resultat: 87 noder (65 substantiver + 22 hjemmesider), 243 forbindelser

Tips

Søgning: - Start med 10 søgeord - Brug 10 resultater pr. søgning - Tjek at du får forskellige domæner

Scraping: - Brug altid Level 1 først - Ekskluder sociale medier (LinkedIn, Facebook) - Vær tålmodig - det tager tid (1-5 min pr. 10 URLs)

Netværk: - Brug minimum degree 3-5 for renere netværk - Sammenlign søge- og indholdsnetværk

Problemløsning

Scraping tager for lang tid? - Tjek at du bruger Level 1 (ikke 2 eller 3) - Nogle sider er langsomme

Mange fejl i scraping? - CAPTCHA udfordringer er normale - Systemet bruger search snippets som fallback - Ekskluder problematiske domæner

Netværk genereres ikke? - Sørg for at scraping er helt færdig - Tjek at nogle sider blev scraped succesfuldt - Øg minimum word frequency hvis for mange noder

Samlet tid for hele workflow: 30-45 minutter **Anbefalet antal søgeord:** 4-6 **Anbefalet resultater pr. søgning:** 10