Social Data Science: Potential and challenges

Morten Axel Pedersen (SODAS & Dept. Anthr.)
Andreas Bjerre-Nielsen (SODAS & Dept. Econ.)

KØBENHAVNS UNIVERSITET

SODAS: Copenhagen Center for Social Data Science

It all started with 7500 Dollars in seed money

"Our aim is to establish an interdisciplinary research Center for Computational Social Science at the Faculty of Social Science putting new technological developments to use across the social sciences" (2011)
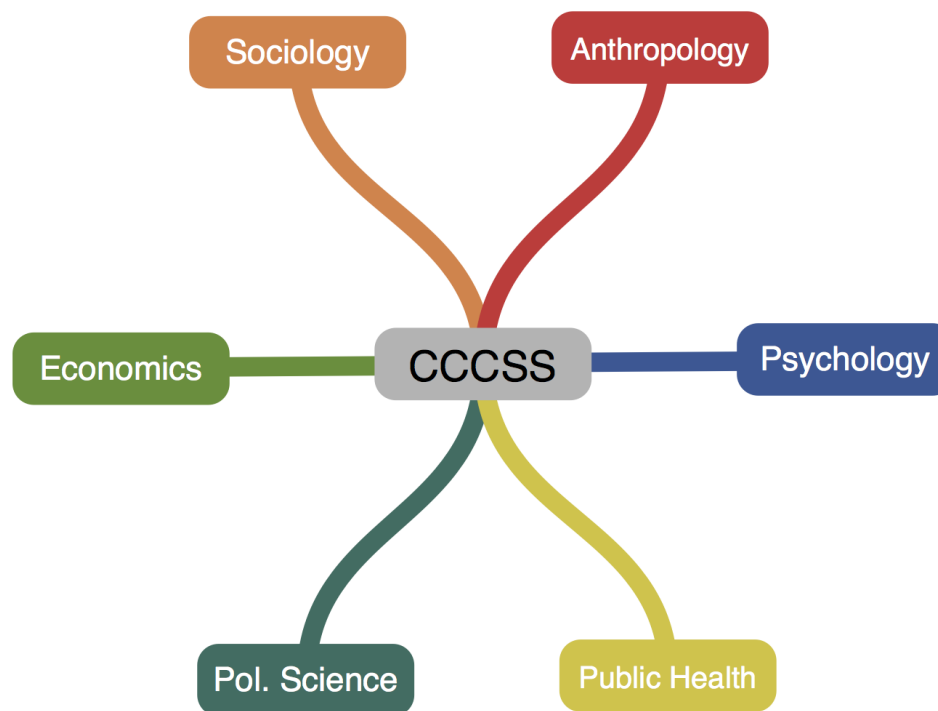
Baggrund: The Copenhagen Network Study

- Map out the complete network (including face-to-face, telecommunication, email, chat, and online social networks).

- Entire freshman class at DTU ($N > 1000$)

- Smart phones running software for data collection.

- Cross-disciplinary approach, including ethnography

number of bits per user

X

number of users

Face-to-face

Texts

Calls

12 pm - 6 am   6 -12 am   12 am - 6 pm   6 - 12 pm

A number of Facebook friends

B volume of interaction with Facebook friends

C number of voice call contacts

D number of SMS contacts

E Pearson correlation between BFI personality traits and communication channels

Personality trait	N_{fs}	N_{ff}	N_c	N_s
Openness	—	—	—	—
Conscientiousness	—	0.10*	—	—
Extraversion	0.28***	0.25***	0.24***	0.25***
Agreeableness	—	—	0.09*	—
Neuroticism	—	—	—	—
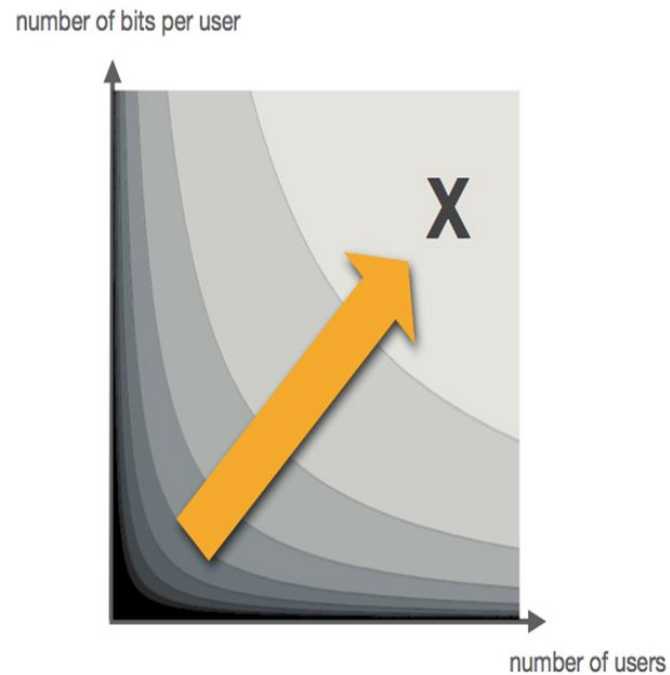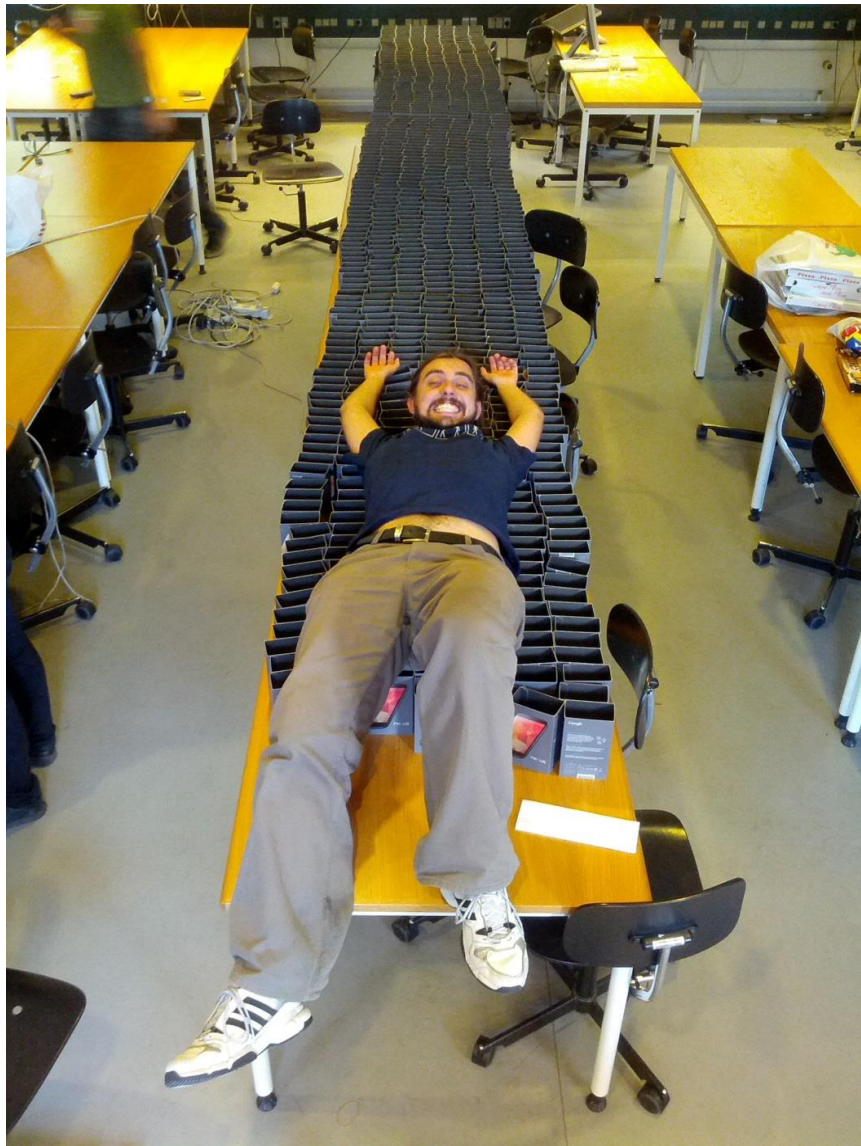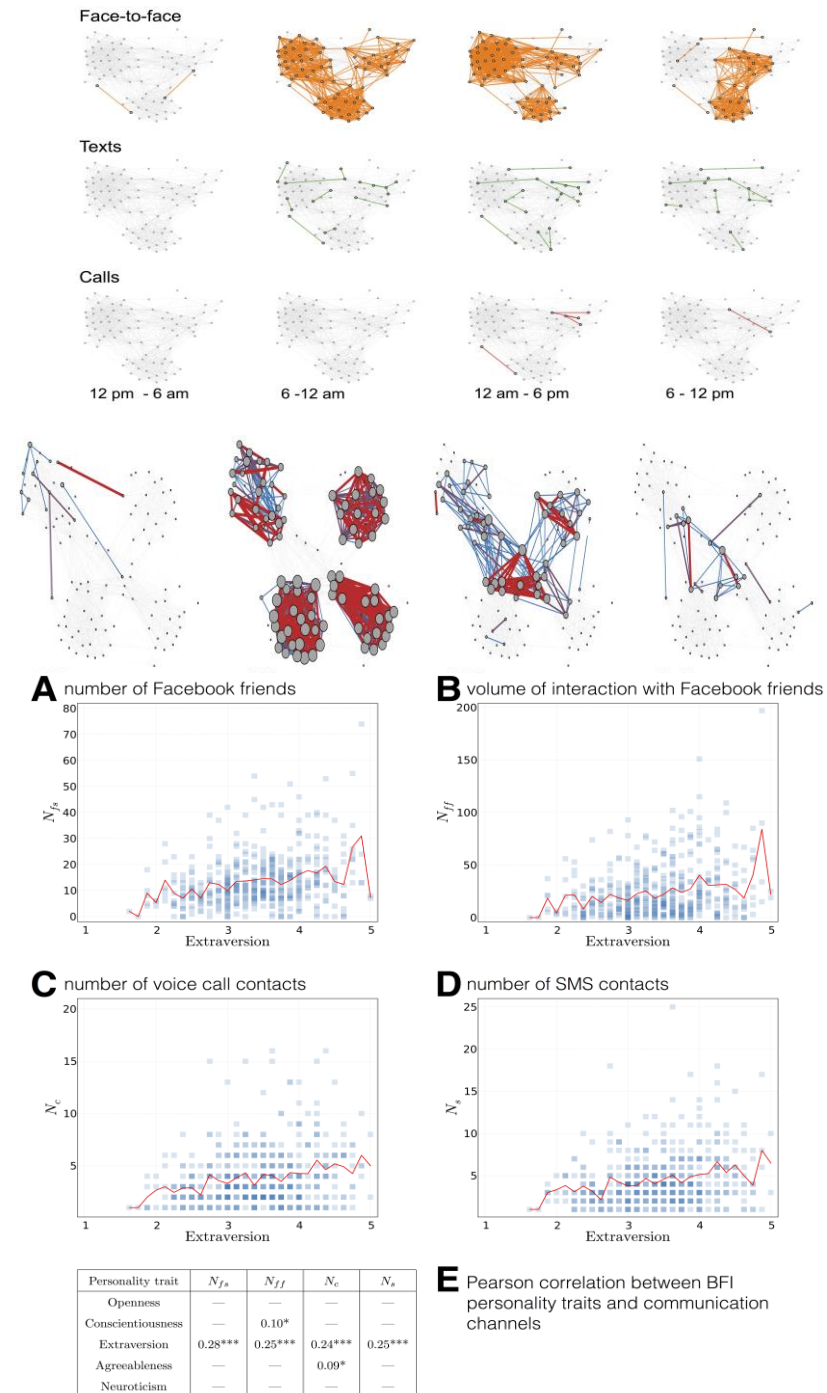
Fundamental structures of dynamic social networks

Vedran Sekara, Arkadiusz Stopczynski, and Sune Lehmann

Significance

We study the dynamic network of real world person-to-person interactions between approximately 1,000 individuals with 5-min resolution across several months. There is currently no coherent theoretical framework for summarizing the tens of thousands of interactions per day in this complex network, but here we show that at the right temporal resolution, social groups can be identified directly. We outline and validate a framework that enables us to study the statistical properties of individual social events as well as series of meetings across weeks and months. Representing the dynamic network as sequences of such meetings reduces the complexity of the system dramatically. We illustrate the usefulness of the framework by investigating the predictability of human social activity.
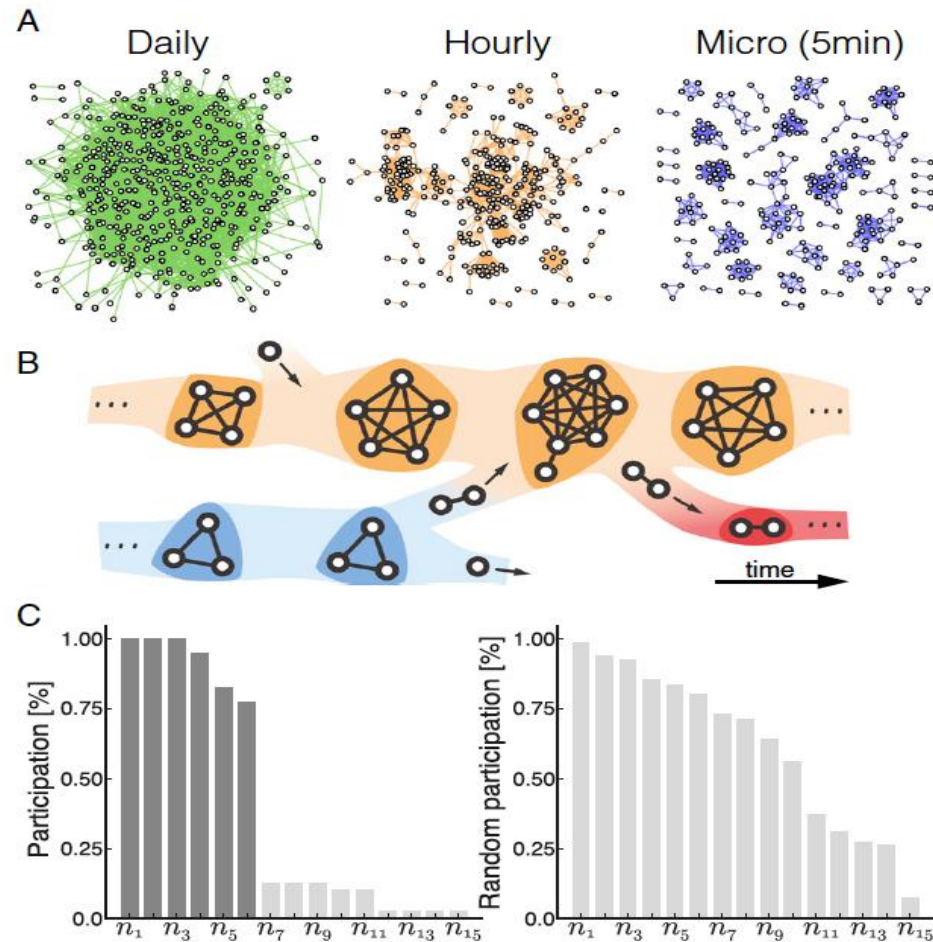
Fig. 1. Properties of gatherings. (*A*) The social network at different timescales. The network formed by face-to-face meetings within 1-d (green), 60-min (orange), and 5-min (blue) temporal aggregation. In the 5-min time slices, groups are directly observable without much ambiguity, but the overlap between groups increases as time is aggregated across bins. (*B*) Illustration of gathering dynamics. Gatherings evolve gradually with members flowing in and out of social contexts. (*C*) Extracting cores from participation profiles. Dark-gray bars denote nodes with participation levels above the maximal gap. Ordered participation profile for empirical data (*Left*), as a null model we use participation profiles generated from a uniform random distribution (*Right*).

BIG DATA & SOCIETY

Stitching together the heterogeneous party: A complementary social data science experiment

Anders Blok, Hjalmar B Carlsen, Tobias B Jørgensen, more...   Show all authors ⌄

First Published November 6, 2017 | Research Article | Check for updates

Download PDF   Article information ⌄   Altmetric 5

Abstract

The era of 'big data' studies and computational social science has recently given rise to a number of realignments within and beyond the social sciences, where otherwise distinct data formats – digital, numerical, ethnographic, visual, etc. – rub off and emerge from one another in new ways. This article chronicles the collaboration between a team of anthropologists and sociologists, who worked together for one week in an experimental attempt to combine 'big' transactional and 'small' ethnographic data formats. Our collaboration is part of a larger cross-disciplinary project carried out at the Danish Technical University (DTU), where high-resolution transactional data from smartphones allows for recordings of social networks amongst a freshman class (N = 800). With a parallel deployment of ethnographic fieldwork among the DTU students, this research set-up raises a number of questions concerning how to assemble disparate 'data-worlds' and to what epistemological and political effects? To address these questions, a specific social event – a lively student party – was singled out from the broader DTU dataset. Our experimental collaboration used recordings of Bluetooth signals between students' phones to visualize the ebb and flow of social intensities at the DTU party, juxtaposing these with ethnographic field-notes on shifting party atmospheres. Tracing and reflecting on the process of combining heterogeneous data, the article offers a concrete case of how a 'stitching together' of digital and ethnographic data-worlds might take place.

Keywords

Big social data, ethnography, data-worlds, experiment, complementarity, party sociality

---

29. januar, 2019 — 08:36

VIDENSKAB

FORSKERE VIL FINDE OPSKRIFTEN PÅ DEN GODE FEST

Arkivfoto fra fest på CSS Image: Lars Juul Hauschildt

Big data — Antropologer kombinerer big data og feltnoter for at kortlægge den gode fest. I fremtiden håber de på at kunne teste klassiske teorier inden for samfundsvidenskaben ved hjælp af metoden.

A POLITICAL ATTENTION EXPERIMENT

A social data science study of the dynamics
of attention during a major political - or
indeed a critical epidemic - event

- "Extended case study" 2.0

- Open-ended pilot

- Hypothesis-testing/intervention

- Social media data

- Ethnographic data

- Infrastructure data

- Wifi data

UNIVERSITY OF COPENHAGEN

"Social Fabric" / "Sensible DTU"

- Map out the complete network (including face-to-face, telecommunication, email, chat, and online social networks).
- Entire freshman class at DTU ($N > 1000$)
- Smart phones running software for data collection.
- Cross-disciplinary approach, including ethnography

number of bits per user

X

number of users

Videnskab·dk
Danmarks førende videnskabsmedie

Temaer   For skoler   Nyhedsbreve   Center for Faglig Formidling   Om Videnskab.dk

Nyheder   COVID-19   Red Verden   Krop & Sundhed   Kultur & Samfund   Naturvidenskab   Teknologi & Innovation

Corona-debatten på Twitter har taget en kraftig drejning efter nedlukningen af Danmark – men hvorfor?

Før landet blev lukket ned, var Twitter præget af corona-frygt, men nu har den kritiske samtale taget over, viser ny undersøgelse

Weekendavisen | Ideer   SEKTIONER   #30 24. JULI 2020

Coronaliv. På et københavnsk kollegie har coronanedlukningen ikke ført til ensomhed, men til et stærkere bånd til de nærmeste. Ny antropologisk undersøgelse viser positive aspekter ved krisen.

Farvel til smalltalk

▶ LYT TIL ARTIKLEN

KÅRE HOLM THOMSEN

Da coronakrisen lukkede Danmark ned og tvang folk til at holde afstand fra hinanden, førte det til ensomhed og socialt afsavn. Sådan har meldingen lydt fra både forskere og interesseorganisationer de seneste måneder. Alzheimerforeningen og Ældre Sagen har sagt, at deres medlemmer har følt sig ensomme i krisen, og også de unge er psykisk mærket, lyder det: Dansk Magisterforening fremhæver for eksempel, at andelen af de studerende på landets universiteter, som føler sig ensomme på deres studier, er steget.

Men måske er frygten for ensomheden og dens konsekvenser overdrevet. I et nyt antropologisk studie har forskere fulgt en gruppe unge tæt under coronanedlukningen, og deres observationer viser, at de unge var overraskende gode til at tilpasse sig den nye virkelighed.

»De unge, vi har undersøgt, rapporterer ikke udtalte grader af socialt afsavn,« siger antropolog Kristoffer Albris, adjunkt på Institut for Antropologi og Copenhagen Center for Social Data Science ved Københavns Universitet, der er en af forskerne bag studiet.

COVID-19 projects

List of ongoing research projects with relation to the COVID-19 cricis with SODAS people as PI or partners.

COVID-19 Snapshot MOnitoring in Denmark (COSMO Denmark)
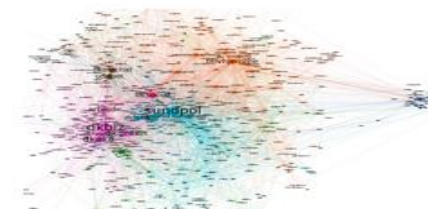
Cross-national dynamics of risk related attitudes on social media during the COVID-19 outbreak

HOPE - How Democracies Cope with Covid19: A Data-Driven Approach

Solidarity and Volounteering in the Corona Crisis

The Dynamics of Political Discourse and Attention during the COVID-19 outbreak

What do the Danes think, know and mean about the Corona Epidemic?

Changing labour market/demand for new skills

Could a Computational Social Scientist be Your Next Best Hire?

Social Scientists are experts are using data to find insights on behavior and relationships between groups of people.

NY UDDANNELSE PÅ SAMF

Kandidat i social datavidenskab

JOB-TREND

Arbejdsgiverne efterspørgsel på folk der kan kombinere samfundsvidenskab og datavidenskab, stiger støt, viser analyse fra forskerne på SODAS
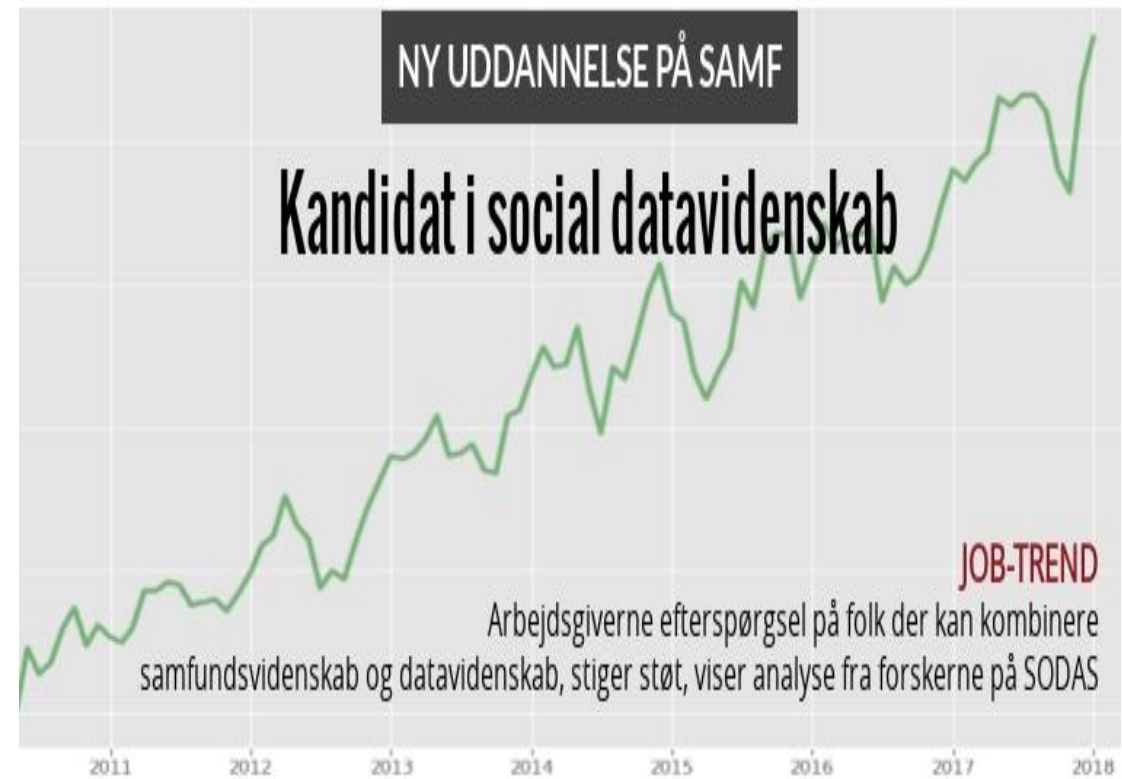
Master's programmes
Guide to studies and admission

Search... All 🔍

■ MSc in Social Data Science

❯ Admission requirements
❯ Application procedure
❯ Profile and career
❯ Programme structure
❯ Meet the teachers
❯ Student life
❯ FAQ
❯ Contact

Related Information

❯ Master's Programmes

Master of Science (MSc) in Social Data Science

PROFILE AND CAREER ❯

PROGRAMME STRUCTURE ❯

ADMISSION REQUIREMENTS ❯

Related Programmes

- Anthropology
- Computer Science
- Economics
- Global Development
- IT and Cognition
- Political Science
- Psychology (in Danish)

New types of data, in particular digital data, are flooding individual lives, organisations and societies all over the world. The broad catchphrase for the analysis of such digital traces is 'data science'. In our master's degree programme, we integrate the analysis of such large-scale and unstructured data – 'big data' – with social scientific modes of enquiry. We call this integration social data science (SDS).

The master's degree programme combines state-of-the-art data science techniques, including machine learning, predictive analytics and natural language processing with core sociological, economic, anthropological,
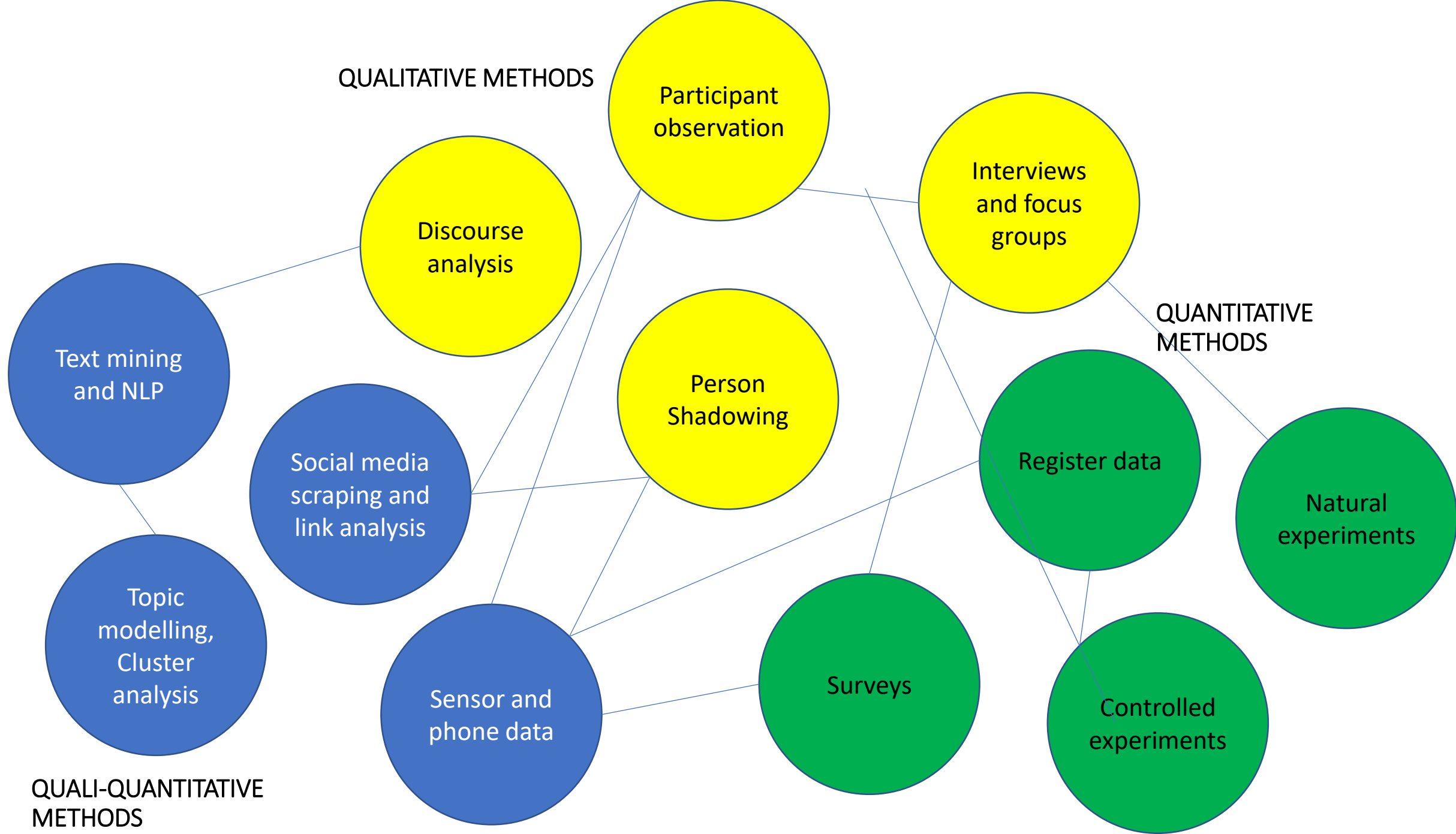
QUALITATIVE METHODS

Participant observation

Interviews and focus groups

Discourse analysis

QUANTITATIVE METHODS

Text mining and NLP

Person Shadowing

Register data

Natural experiments

Social media scraping and link analysis

Topic modelling, Cluster analysis

Sensor and phone data

Surveys

Controlled experiments

QUALI-QUANTITATIVE METHODS

Do..

1. Design a shared 'machine"/experiment in the middle

2. Foster a culture of permitted failure and psych. safety

3. Facilitate intense collaborative moments (data-sprints)

4. Secure a physical space for core team and events

5. Support outreach (key platform for interdisciplinarity)

Don't..

1. Submit to a Jane Austen fantasy of romantic love

2. Impose a single methodology/theory/data-standard

3. Only go for top journals and reproduce hierarchies

4. Accept 'innocent' jokes about minor sciences/juniors

5. Assume that you can remain faithful to your discipline
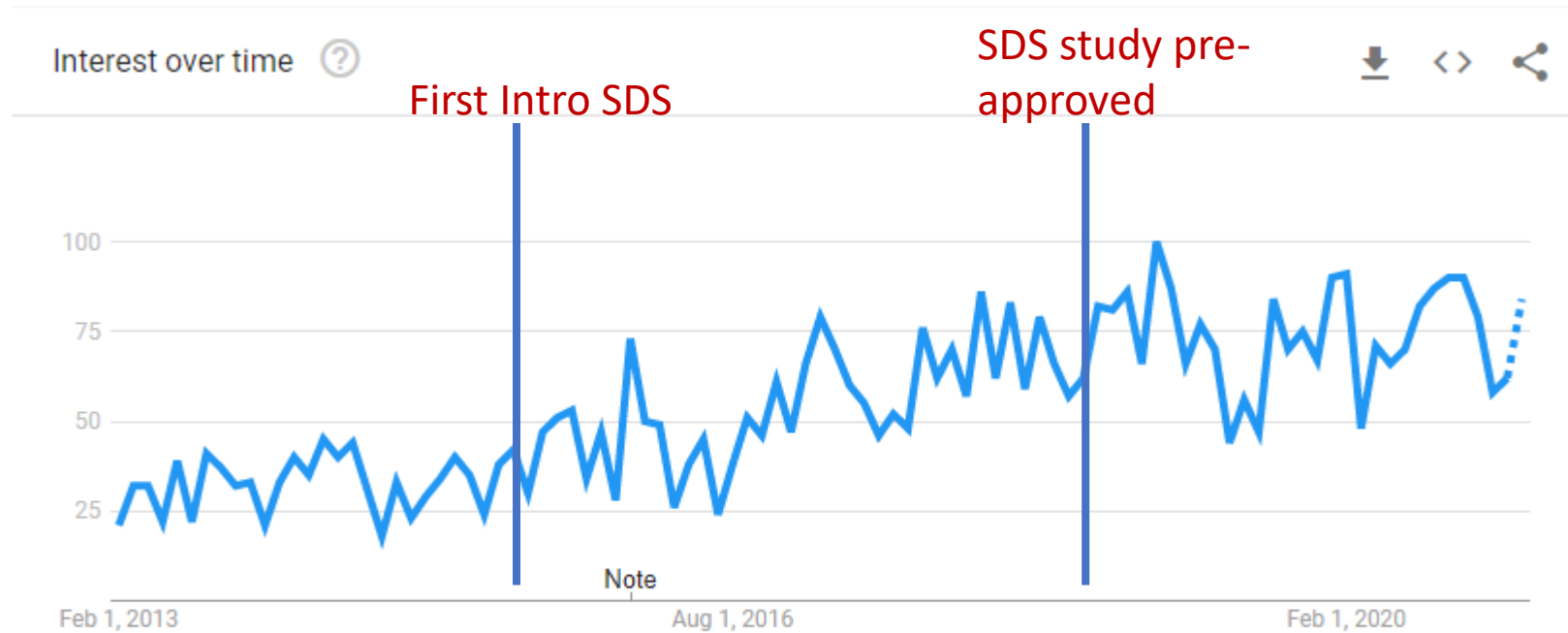
How to start a new degree?

Background: start to teach courses

- Seminar in machine learning and network science (2013)

- "Social Data Science" course (2015)
  - Data science / machine learning for social science (e.g. poli. sci, sociology)
  - Grew from ~50 enrolled in 2017 to ~220 in 2020!

Higher ambitions

➢ advanced courses (2017 approval, start 2018)

➢ degree program

Social Data Science – the journey

Great artists steal ..

Approval of study in 2018/19.

- Top UK universities quick to follow suit:

Challenges in opening program

- Accepted study program was "rough design"
  - Scaling from few courses to degree program
  - Each scholar with own vision of "optimal program"

- Procedure: Put experts of their fields together in meetings:
  - Shared vision of what field consists of:
    - E.g. 20% qualitative/explorative
  - Leverage "wisdom of the crowd" and find consensus

- Collective refinement helped ensure
  - Skill progression was feasible
  - Most relevant skills are taught
    - E.g. machine learning/statistics without linear algebra

Program content

	The master's programme in Social Data Science (120 ECTS)			
	Block 1		*Block 2*	
Semester 1 (Autumn, year 1)	Social Data Science Base Camp (15 ECTS)			
	Elementary Social Data Science (7.5 ECTS)		Social Data Analysis (7.5 ECTS)	
	Block 3		*Block 4*	
Semester 2 (Spring, year 1)	Advanced Social Data Science I (7.5 ECTS)	Data Governance: Law, Ethics, and Politics (7.5 ECTS)	Advanced Social Data Science II (7.5 ECTS)	Digital Methods (7.5 ECTS)
Semester 3 (Autumn, year 2)	Elective courses and mobility window (30 ECTS)			
Semester 4 (Spring, year 2)	Master's Thesis (30 ECTS)			

Program content

Introduction to Programming in Python

	The master's programme in Social Data Science (120 ECTS)	
	Block 1	*Block 2*
Semester 1 (Autumn, year 1)	Social Data Science Base Camp (15 ECTS)	
	Elementary Social Data Science (7.5 ECTS)	Social Data Analysis (7.5 ECTS)
	Block 3	*Block 4*
Semester 2 (Spring, year 1)	Advanced Social Data Science I (7.5 ECTS) — Data Governance: Law, Ethics, and Politics (7.5 ECTS)	Advanced Social Data Science II (7.5 ECTS) — Digital Methods (7.5 ECTS)
Semester 3 (Autumn, year 2)	Elective courses and mobility window (30 ECTS)	
Semester 4 (Spring, year 2)	Master's Thesis (30 ECTS)	

Program content

Data collection:
scraping, survey,
experiments,
combination
with ethnographic

	The master's programme in Social Data Science (120 ECTS)			
	Block 1		*Block 2*	
Semester 1 (Autumn, year 1)	Social Data Science Base Camp (15 ECTS)			
	Elementary Social Data Science (7.5 ECTS)		Social Data Analysis (7.5 ECTS)	
	Block 3		*Block 4*	
Semester 2 (Spring, year 1)	Advanced Social Data Science I (7.5 ECTS)	Data Governance: Law, Ethics, and Politics (7.5 ECTS)	Advanced Social Data Science II (7.5 ECTS)	Digital Methods (7.5 ECTS)
Semester 3 (Autumn, year 2)	Elective courses and mobility window (30 ECTS)			
Semester 4 (Spring, year 2)	Master's Thesis (30 ECTS)			

Program content

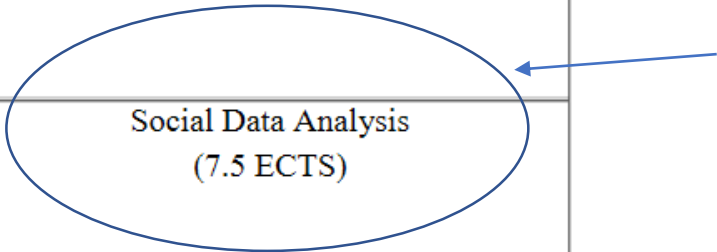|	The master's programme in Social Data Science (120 ECTS)			
	Block 1		*Block 2*	
Semester 1 (Autumn, year 1)	Social Data Science Base Camp (15 ECTS)			
	Elementary Social Data Science (7.5 ECTS)		Social Data Analysis (7.5 ECTS)	
	Block 3		*Block 4*	
Semester 2 (Spring, year 1)	Advanced Social Data Science I (7.5 ECTS)	Data Governance: Law, Ethics, and Politics (7.5 ECTS)	Advanced Social Data Science II (7.5 ECTS)	Digital Methods (7.5 ECTS)
Semester 3 (Autumn, year 2)	Elective courses and mobility window (30 ECTS)			
Semester 4 (Spring, year 2)	Master's Thesis (30 ECTS)			

Quantitative methods: asic statistics, supervised/ unsupervised machine learning, quasi-experimental methods, network analysis

Program content

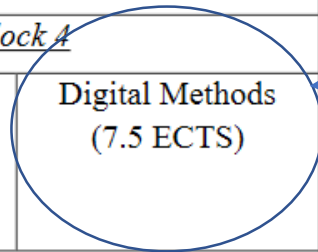|	The master's programme in Social Data Science (120 ECTS)			
	Block 1		*Block 2*	
Semester 1 (Autumn, year 1)	Social Data Science Base Camp (15 ECTS)			
	Elementary Social Data Science (7.5 ECTS)		Social Data Analysis (7.5 ECTS)	
	Block 3		*Block 4*	
Semester 2 (Spring, year 1)	Advanced Social Data Science I (7.5 ECTS)	Data Governance: Law, Ethics, and Politics (7.5 ECTS)	Advanced Social Data Science II (7.5 ECTS)	Digital Methods (7.5 ECTS)
Semester 3 (Autumn, year 2)	Elective courses and mobility window (30 ECTS)			
Semester 4 (Spring, year 2)	Master's Thesis (30 ECTS)			

Core theory: behavior, networks and ideas

Program content

	The master's programme in Social Data Science (120 ECTS)			
	Block 1		*Block 2*	
Semester 1 (Autumn, year 1)	Social Data Science Base Camp (15 ECTS)			
	Elementary Social Data Science (7.5 ECTS)		Social Data Analysis (7.5 ECTS)	
	Block 3		*Block 4*	
Semester 2 (Spring, year 1)	Advanced Social Data Science I (7.5 ECTS)	Data Governance: Law, Ethics, and Politics (7.5 ECTS)	Advanced Social Data Science II (7.5 ECTS)	Digital Methods (7.5 ECTS)
Semester 3 (Autumn, year 2)	Elective courses and mobility window (30 ECTS)			
Semester 4 (Spring, year 2)	Master's Thesis (30 ECTS)			

Program content

	The master's programme in Social Data Science (120 ECTS)			
	Block 1		*Block 2*	
Semester 1 (Autumn, year 1)	Social Data Science Base Camp (15 ECTS)			
	Elementary Social Data Science (7.5 ECTS)		Social Data Analysis (7.5 ECTS)	
	Block 3		*Block 4*	
Semester 2 (Spring, year 1)	Advanced Social Data Science I (7.5 ECTS)	Data Governance: Law, Ethics, and Politics (7.5 ECTS)	Advanced Social Data Science II (7.5 ECTS)	Digital Methods (7.5 ECTS)
Semester 3 (Autumn, year 2)	Elective courses and mobility window (30 ECTS)			
Semester 4 (Spring, year 2)	Master's Thesis (30 ECTS)			

Integrating "quantitative and qualitative" analysis

Program content

	The master's programme in Social Data Science (120 ECTS)			
	Block 1		*Block 2*	
Semester 1 (Autumn, year 1)	Social Data Science Base Camp (15 ECTS)			
	Elementary Social Data Science (7.5 ECTS)		Social Data Analysis (7.5 ECTS)	
	Block 3		*Block 4*	
Semester 2 (Spring, year 1)	Advanced Social Data Science I (7.5 ECTS)	Data Governance: Law, Ethics, and Politics (7.5 ECTS)	Advanced Social Data Science II (7.5 ECTS)	Digital Methods (7.5 ECTS)
Semester 3 (Autumn, year 2)	Elective courses and mobility window (30 ECTS)			
Semester 4 (Spring, year 2)	Master's Thesis (30 ECTS)			

Electives, internship, data-collection, Exchange program

Challenges in interdisciplinary program

- Differences in background skills
  - Frustration with teaching level
  - Some due to difficulty, others from boredom

  - Focus on students without previous quant skills
    - Communication: encouragement, extra tutoring
    - Leverage strong group work to learn

  - Focus on students with previous quant skills
    - Communication: remind improved learning etc.
    - Open ended exam projects, bonus exercises
    - Potential electives for basic courses?

A Rosetta Stone for students

- How to bring people from different backgrounds together?
  - Problem: divide between "what is knowledge"

- Introduction week
  - Mix into different groups

  - Share experiences of working with qualitative and quantitative data

  - E.g. field work in supermarkets – interview and observation