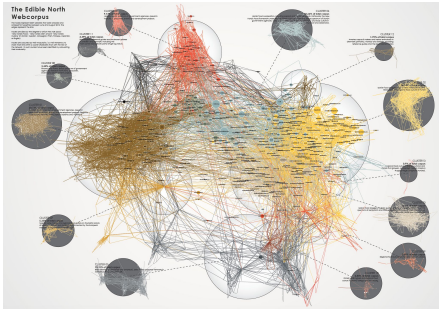
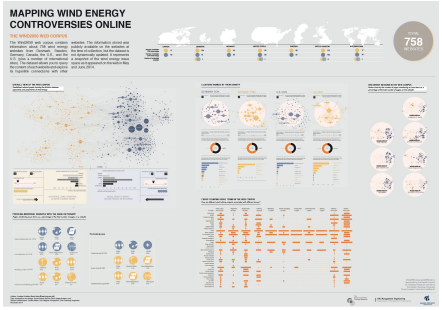
From data science as anthropology to anthropology as data science

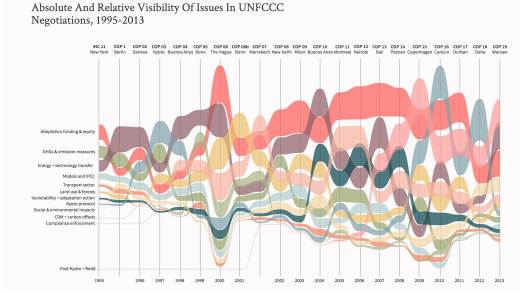
Anders Kristian Munk

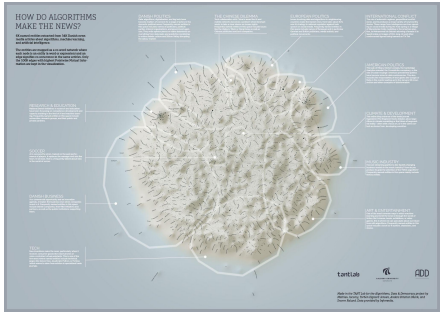
@AndersKMunk / @TANTlab
www.tantlab.aau.dk

tantlab









Studying science and technology in society

Particularly:

- controversies about new science and technology
- and often about datafication, machine learning, AI, algorithms etc.
- as they play out on digital media

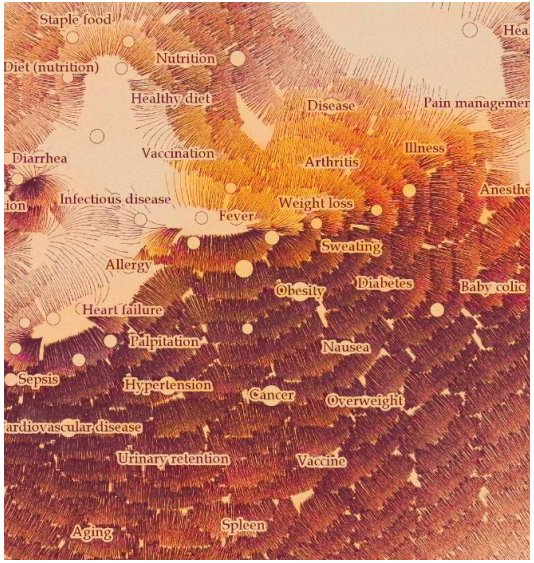
With anthropological and computational methods

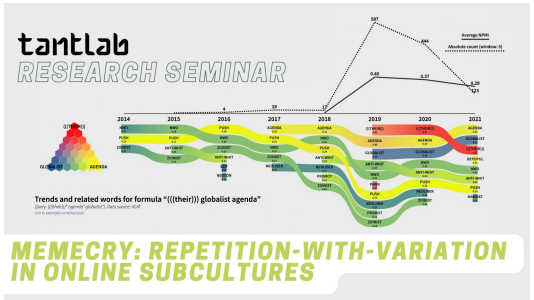
Combining:

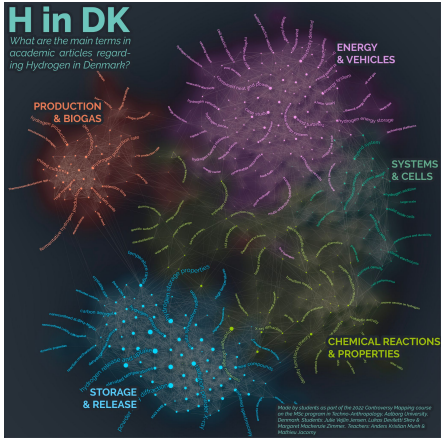
- Ethnography
- Big sociocultural data
- Machine learning
- Natural Language Processing
- And exploratory data viz

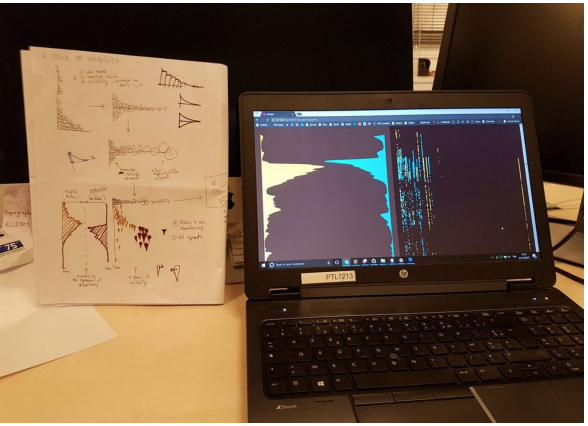
And thus developing:

- New computational SSH research practices
- Plus the tools and technical competencies to support them









In collaboration with partners and stakeholders

As participatory data design

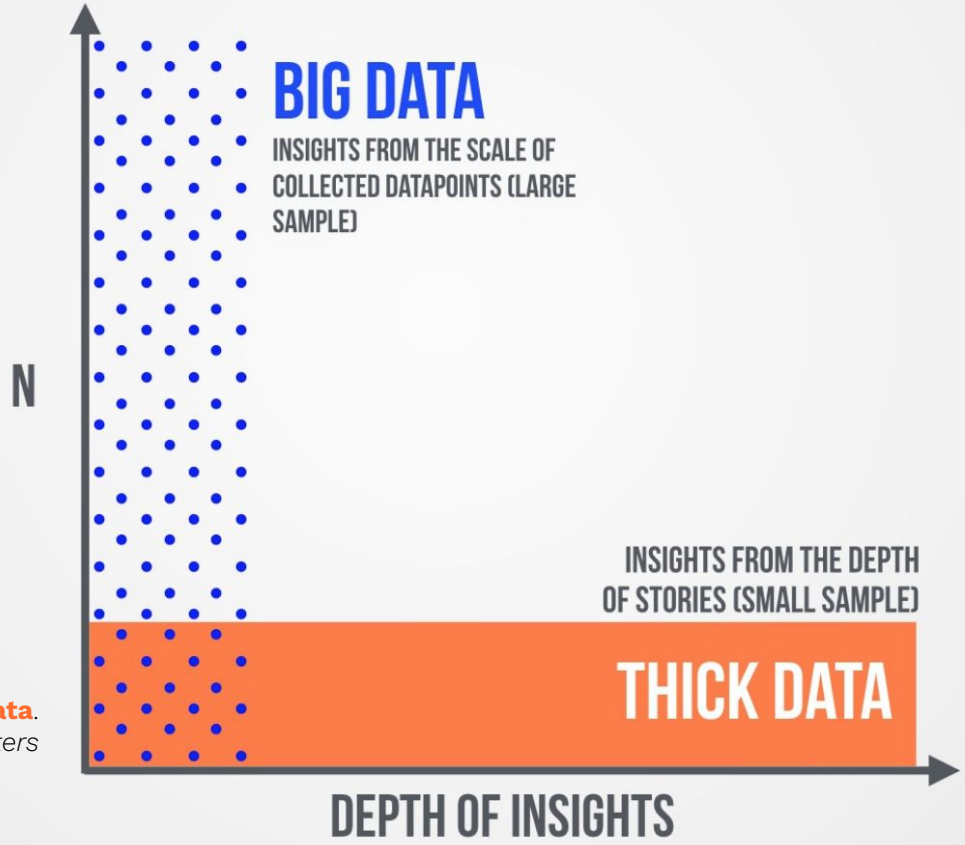
Through sprints and similar formats

To engage publics

Ensure democratic robustness

And build capacity

TO FORM A COMPLETE PICTURE, BOTH BIG AND THICK DATA ARE CRITICAL BECAUSE THEY PRODUCE DIFFERENT TYPES OF INSIGHTS AT VARYING SCALES AND DEPTHS

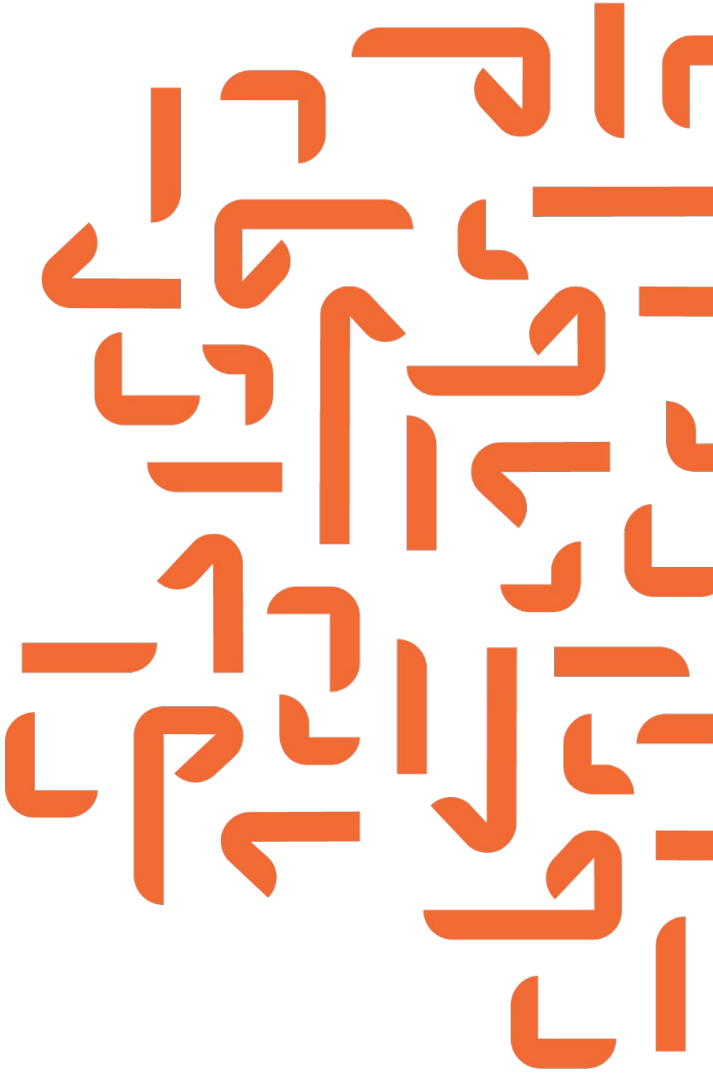


Wang, T. (2013). **Big data needs thick data**. *Ethnography matters*

Beyond big vs. thick

"**Ethnography needs to create a new paradigm** that shifts from shaping its identity by claiming to understand the 'whys' around cultural behaviour to one where it and **Big Data are actually focusing on the same epistemological field**, one that is situated within human behaviour and crucially cultural interpretation. **Both should therefore be seen as interpretive epistemological approaches to analysis for human behaviour and cultural dispositions**."

Curran, J. (2013). **Big data or 'big ethnographic data'? Positioning big data within the ethnographic space**. *Ethnographic praxis in industry conference proceedings*

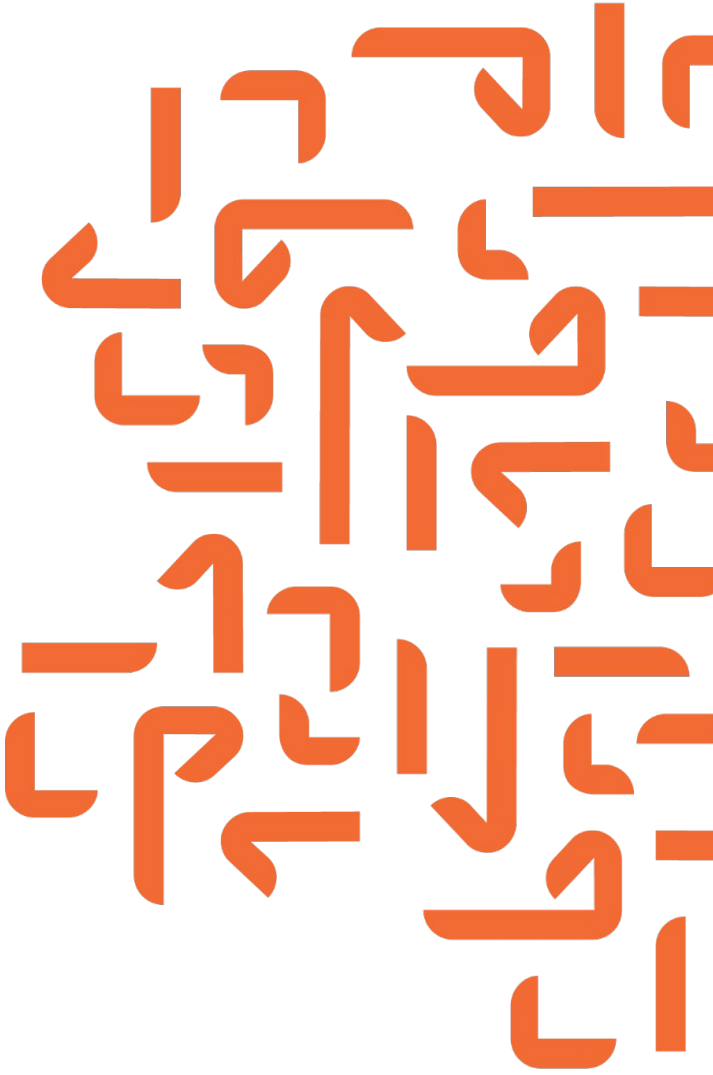


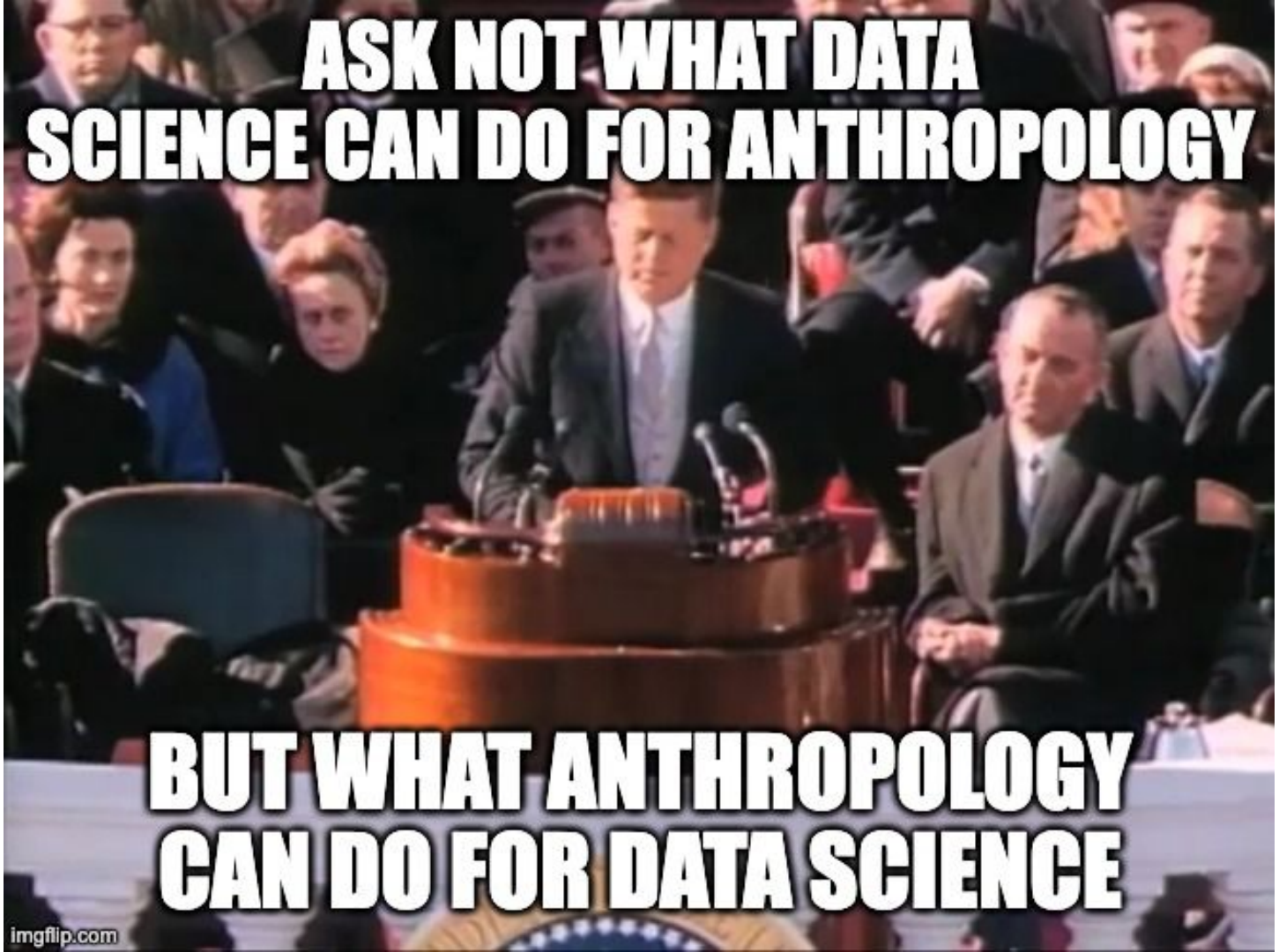
Beyond big vs. thick

“The shifts within machine-learning algorithm development give impetus for incorporating **quantitative techniques that are abductive, local, and interpretive**. The debate between universal and proscriptive versus local and bottom-up knowledge production may not—or at least no longer—need to divide quantitative and qualitative techniques. **Machine learning algorithms “leave open the possibility of situated knowledge production**, entangled with narrative,” a clear parallel to qualitative ethnographic techniques”

Paff, S. (2022). **Anthropology by Data Science.** *Annals of Anthropological Practice*

Quoting **Knox**, H., & **Nafus**, D. (2018). **Introduction: Ethnography for a data-saturated world**. *Ethnography for a data-saturated world*




ASK NOT WHAT DATA
SCIENCE CAN DO FOR ANTHROPOLOGY
BUT WHAT ANTHROPOLOGY
CAN DO FOR DATA SCIENCE
imgflip.com

Mock algebra

"Where W.H.R. Rivers had enthused in Melanesian Society about a future in which kinship studies would "resemble a work on mathematics in which the results will be expressed by symbols, in some cases even in the form of equations," Malinowski was skeptical. "I must frankly confess," he wrote, "that there is not a single account of kinship in which I do not find myself puzzled by some of this **spuriously scientific and stilted mathematization** of kinship facts."

Seaver, N. (2015). **Bastard algebra.** *Data, now bigger and better*

Boyd, J.P. (1969). **The algebra of group kinship.** *Journal of Mathematical Psychology*

(or $XM = Y$). Since $F = F^{-1}$ for the Arunta, the arrow is given two heads, to suggest the symmetry. By convention M-arrows will be solid; and F-arrows dotted since paternity is conjectural.

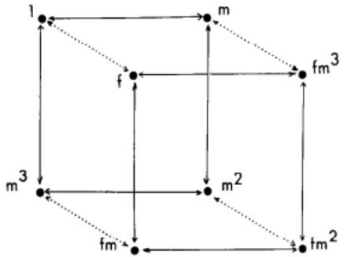


FIG. 1. The graph of the Arunta marriage class system. M-arrows are solid; and F-arrows, dotted.

Since G is regular, we can choose an arbitrary point and label it 1. Then the other points can be labeled with their unique representations $P \in G$. This has also been done in Fig. 1. Table 1 is the "Cayley multiplication table" for the group G of the Arunta tribe.

TABLE 1

THE ARUNTA MULTIPLICATION TABLE[a]

	1	M	M^2	M^3	F	FM	FM^2	FM^3
1	1	M	M^2	M^3	F	FM	FM^2	FM^3
M	M	M^2	M^3	1	FM^3	F	FM	FM^2
M^2	M^2	M^3	1	M	FM^2	FM^3	F	FM
M^3	M^3	1	M	M^2	FM	FM^2	FM^3	F
F	F	FM	FM^2	FM^3	1	M	M^2	M^3
FM	FM	FM^2	FM^3	F	M^3	1	M	M^2
FM^2	FM^2	FM^3	F	FM	M^2	M^3	1	M
FM^3	FM^3	F	FM	FM^2	M	M^2	M^3	1

[a] Row elements are multiplied on the right by column elements. That is, $z = xy$ is the entry of row x, column y.

Since only two of the set $\mathcal{A}$ of basic Arunta kinship terms have been used thus far, we must now investigate how the entire network of relations on E relates to G. Recall

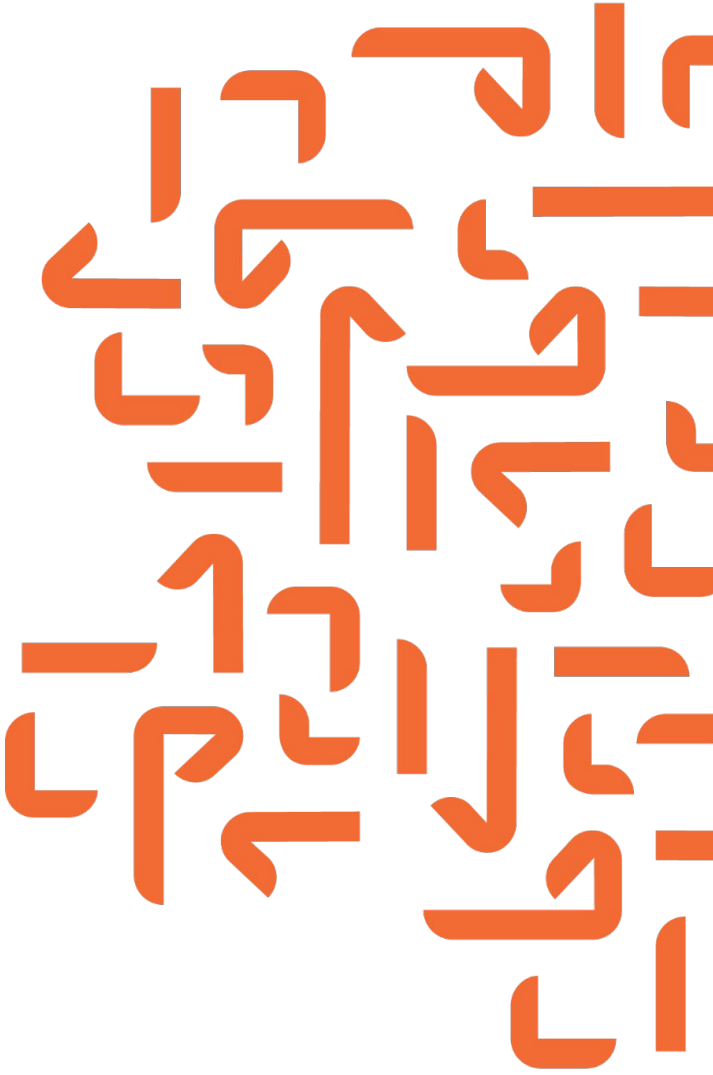
Ethnographic logic

“It is first of all an ***abductive*** logic, taking surprises seriously, and creating new explanations for them. It is also ***iterative***, something that is applied over and over again in the course of a piece of work. And it is ***recursive***, calling on itself to solve a problem that comes up even as it is solving a problem.

(...)

An initial view of X from **POV$_1$**, *changes* through the iterative and recursive application of abductive logic. It changes to a ***different*** view of X that takes into account **contexts** and **meanings** of **POV$_2$**”

Agar, M. (2006). **An Ethnography by Any Other Name...**
Forum for Qualitative Social Research/Sozialforschung

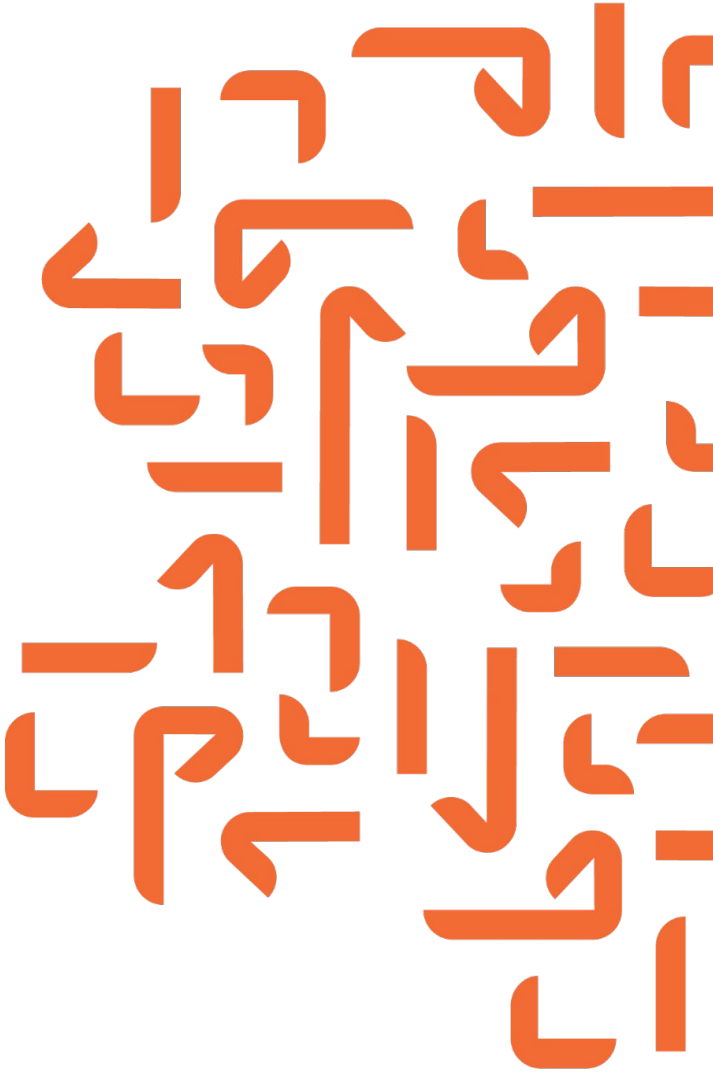


Ethnographic logic

“Once again a major difference between ethnography and the usual social science: **The usual approach**, once it spots something interesting in a particular event, what does it want to do? It wants to ***isolate*** it, figure out how to ***abstract*** it from what’s going on, and then ***measure*** it in some way. That way a larger number of events can be compared.

But when an **ethnographer** focuses on something in an event, the question isn’t, “How do I isolate and measure that?” The question is the context question summarized above. “What else in the event is the sign connected with?” Look around and see what’s going on. Traditional social science is on the lookout for ***variables***; ethnographers are on the lookout for ***patterns***.”

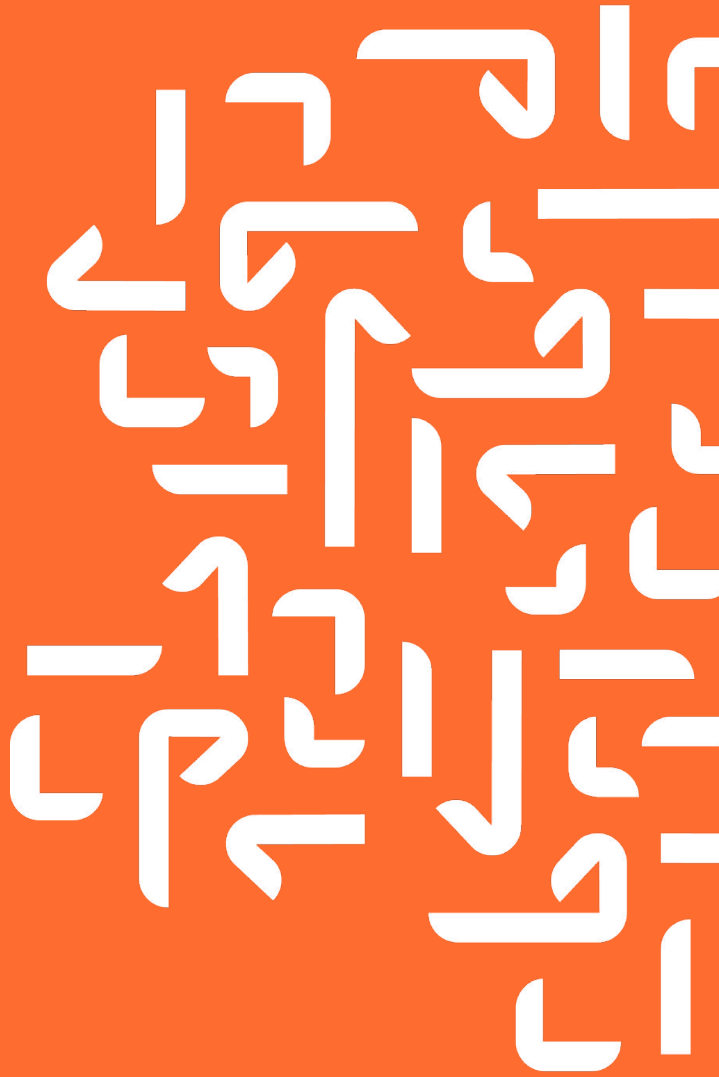
Agar, M. (2006). **An Ethnography by Any Other Name...**
Forum for Qualitative Social Research/Sozialforschung

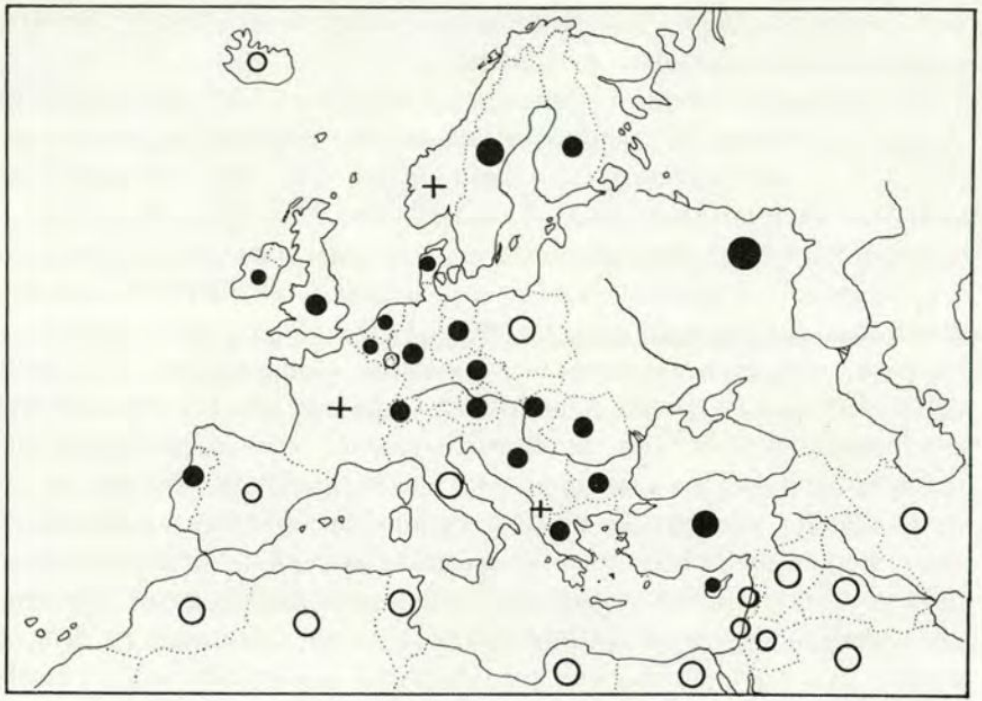


Data science as anthropology

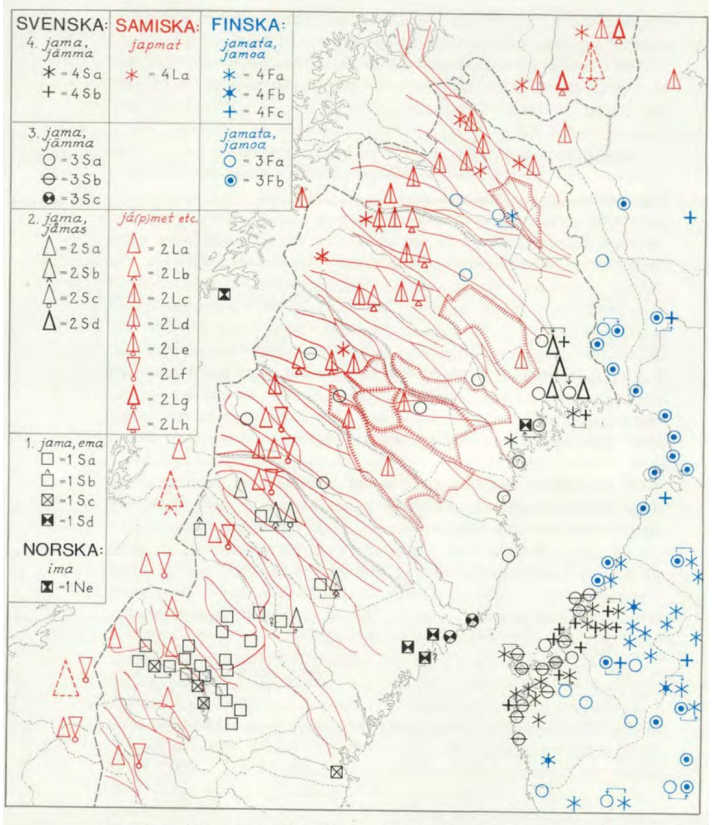
Should help us

- Discover new questions rather than settle the ones we're already asking
- Explore and attune to different POVs rather than eliminate bias
- Provide layers upon layers of interpretation rather than strive for abstract explanations or indeed to be universally explainable.





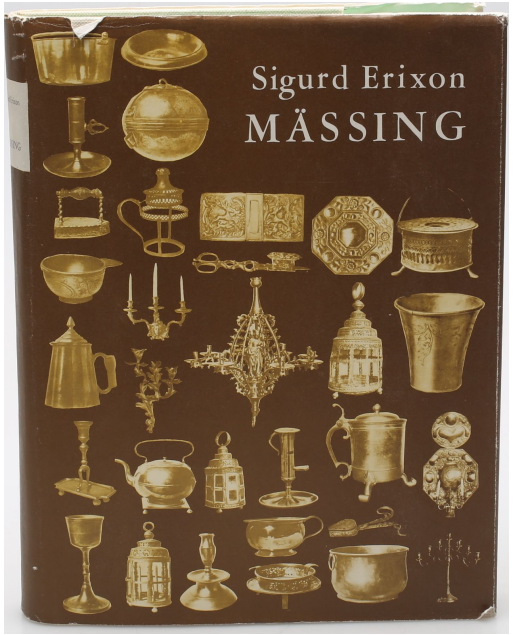
Karta ritad till SIA-konferensen i Zagreb 1966 visande: ● vid konferensen företrädda länder; + skriftligen företrädda länder; ○ andra i planen medtagna länder.

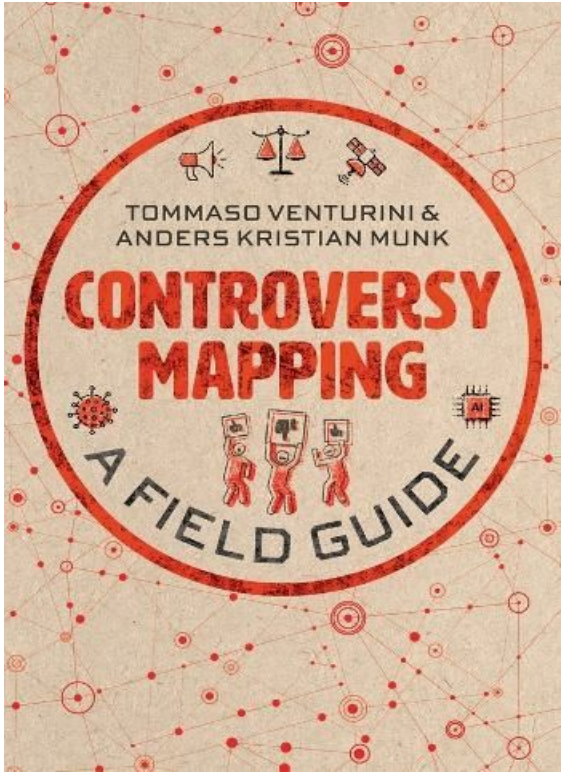


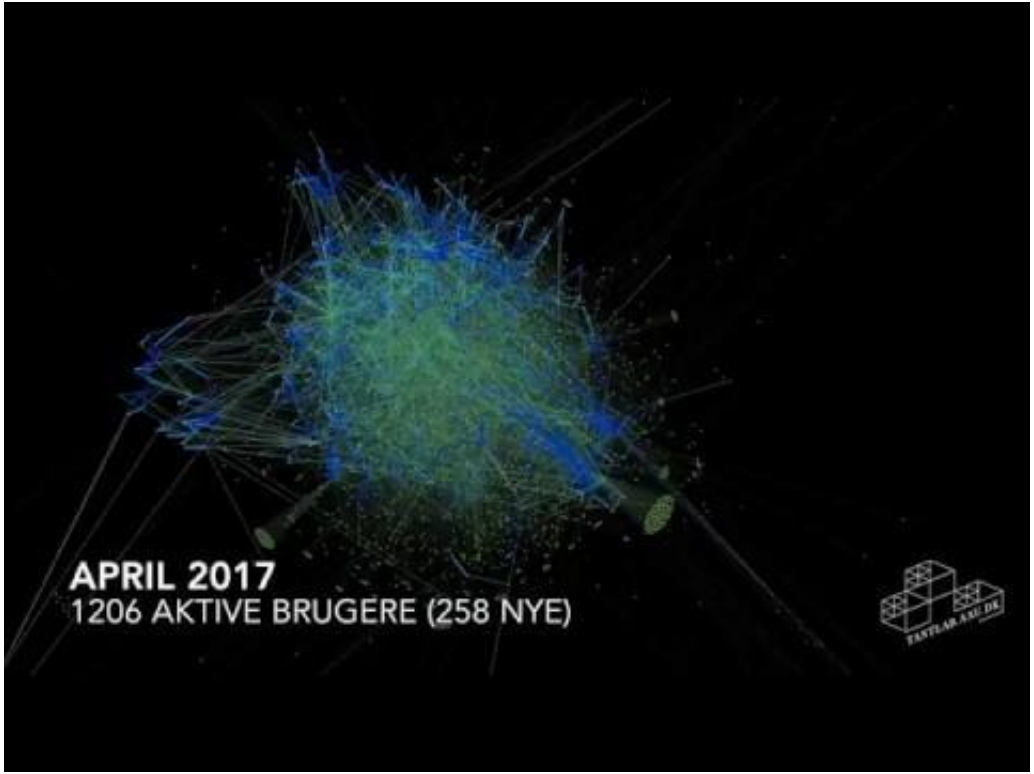
Nyman, Å, (1983). **Internationellt samarbete för etnologisk kartläggning af Europa.**

Between folk cartography and community studies
Massive and coordinated collection + visualisation of cultural ‘traces’ and their spread patterns becomes a vehicle for a uniquely adequate account rather than general laws of culture.

Munk, A. K., **Jensen**, T. E. (2014). **Revisiting the Histories of Mapping: Is there a future for a cartographic ethnology?** *Ethnologia Europaea*







Venturini, T, **Munk**, A.K. (2021). **Controversy Mapping: A Field Guide.** *Polity*

Data people struggle to inspire creative people.

Creative people struggle to convince data people.

Co-funded by the Creative Europe MEDIA Programme of the European Union

Make data more creative.

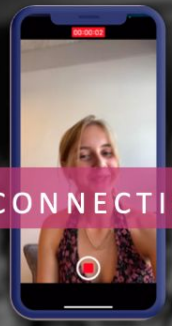
Co-funded by the Creative Europe MEDIA Programme of the European Union

PUBLIKUM / will&agency

Connecting story and audience through AI and anthropology

TREATMENT

EMOTION



CONNECTION

Story, intentions and themes.

Mapping the outside perspective.

Getting the in-depth understanding.

Co-funded by the Creative Europe MEDIA Programme of the European Union

PUBLIKUM / will&agency

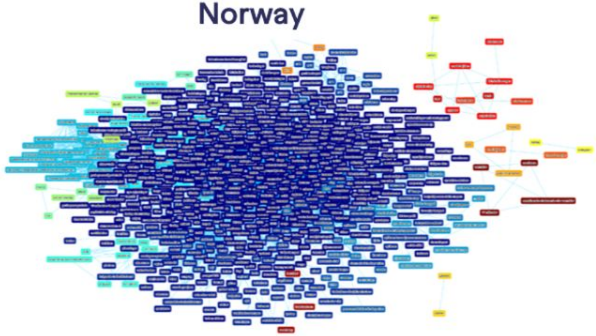
Norsk filminstitutt
OSLO PICTURES
FESTIVAL DE CANNES
UN CERTAIN REGARD
2022 OFFICIAL SELECTION
How is narcissism different in Scandinavia?
SICK OF MYSELF
A film by Kristoffer Borgli
Co-funded by the Creative Europe MEDIA Programme of the European Union
PUBLIKUM / will&agency

Norsk filminstitutt

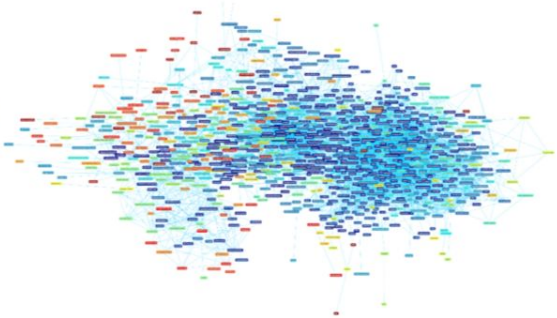
OSLO PICTURES

Narcissism in Scandinavia

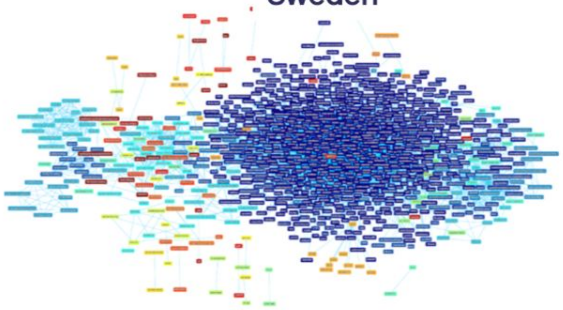
Norway



Denmark



Sweden



Co-funded by the Creative Europe MEDIA Programme of the European Union

PUBLIKUM / will&agency

#3
Archetypes.

#3 Archetypes:

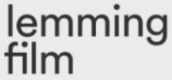
Three versions of a bastard

Denmark: **Dangerous** bastards

Sweden: **Poor** bastard

Norway: **Useful** bastard

Co-funded by the Creative Europe MEDIA Programme of the European Union



Co-funded by the
Creative Europe MEDIA Programme
of the European Union

PUBLIKUM / will&agency

TABLE I. SUBANUN DRINKING TALK

Encounter Stages	Discourse Stages	Focus of Speech Acts	Function
1. Tasting	1. Invitation—permission	Role expression	Assignment of role distances and authority relations to participants
2. Competitive drinking	2. Jar talk	Role expression and context definition	Allocation of encounter resources (turns at drinking and talking)
	3. Discussion 3.1. Gossip 3.2. Deliberation	Topic	Exchange of information; disputation, arbitration; deciding issues on basis of cogent argument
3. Game drinking	4. Display of verbal art	Stylistic	Establishment of euphoria. Deciding issues on basis of skill in use of special styles of discourse (singing, verse)

Segments of a drinking encounter:
1. A turn (continuous drinking by one person)
2. A round (a set of related turns)
3. Encounter stage (a set of related rounds)

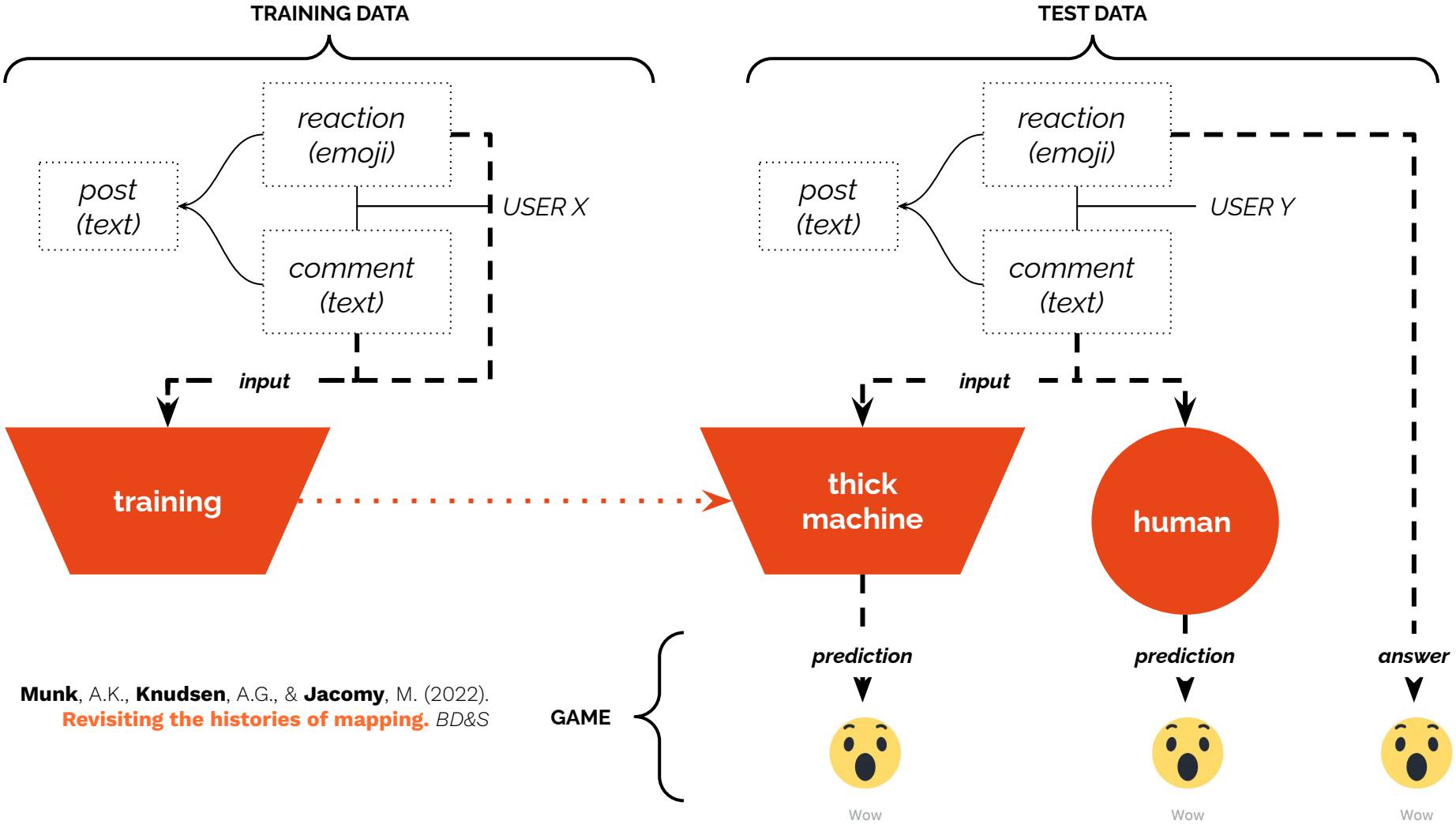
Segments of drinking talk:
1. An utterance (continuous speech by one person)
2. An exchange (a set of related utterances)
3. Discourse stage (a set of related exchanges)

Explainable AI?

"[Ethnoscience] holds that culture is composed of psychological structures by means of which individuals or groups of individuals guide their behavior. [...] from this view of what culture is follows a view, equally assured, of what describing it is—the writing out of systematic rules, an ethnographic algorithm, which, if followed, would make it possible so to operate, to pass (physical appearance aside) for a native."

Geertz, C. (1973). **Thick description**. *The Interpretation of Cultures.*

Frake, C.O. (1964). **How to ask for a drink in Subanun**. *American Anthropologist.*



Munk, A.K., **Knudsen**, A.G., & **Jacomy**, M. (2022).
Revisiting the histories of mapping. *BD&S*

Munk, A.K., **Knudsen**, A.G., & **Jacomy**, M. (2022).
Revisiting the histories of mapping. *BD&S*

POST (translated):

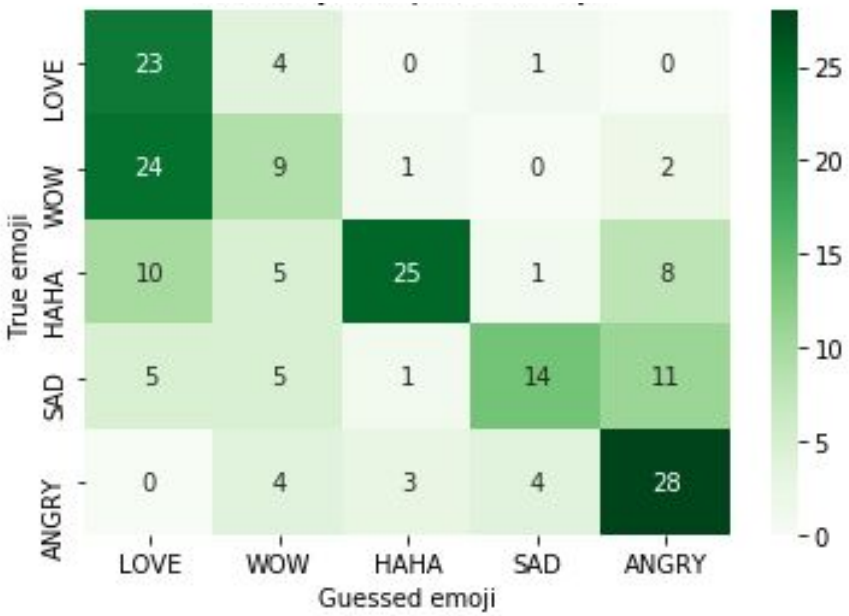
For the Prince, the decision not to be buried next to the Queen is the natural consequence of not receiving the same treatment as his spouse when it comes to the title and function he has always desired, says chief of communications for the Royal House, Lene Balleby.

COMMENT (translated):

I will gladly swap problems with him. Happy to refrain from a royal title in exchange for 8 mio a year. Or just the 29 mio for the sarcophagus. Then he can fight my f***ing ex and the system without getting as much as a penny from me in return. Get a real problem, King Carrot.

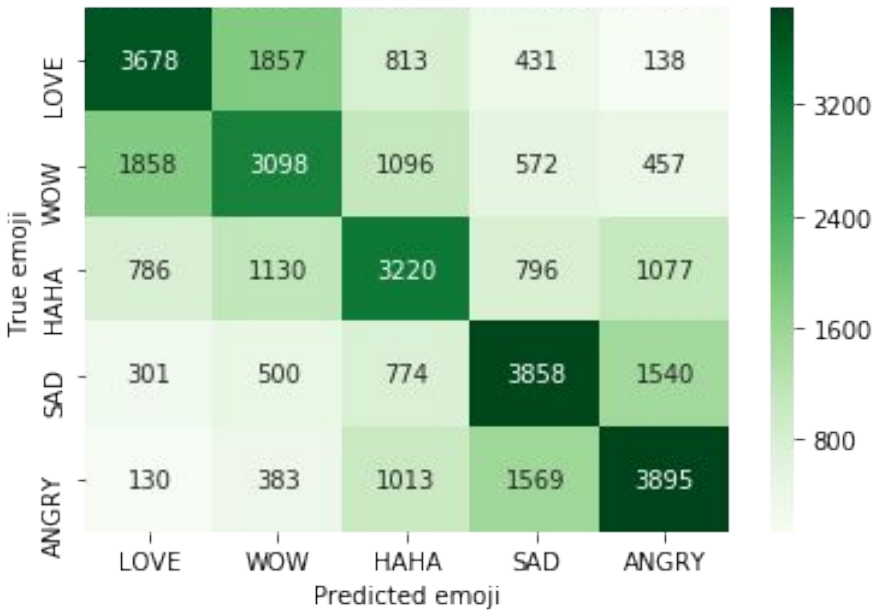
Reaction: 😂
Prediction: 😡

HUMANS OF THE TANT-LAB



Accuracy
52%

NEURAL NETWORK TRAINED ON 140K COMMENTS + REACTIONS

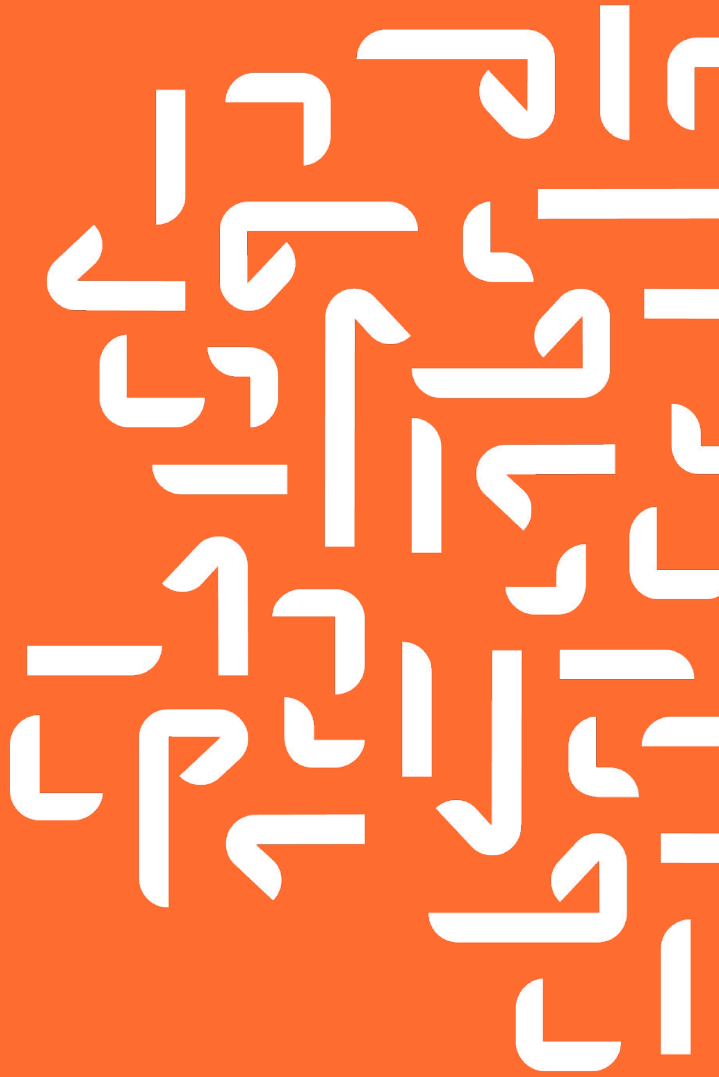


Accuracy
51%

Munk, A.K., **Knudsen**, A.G., & **Jacomy**, M. (2022). **Revisiting the histories of mapping.** *BD&S*

Machine learning is not really (no longer) an explanatory science in search of law

- Is it an interpretative one in search of meaning?
- If so, how should we talk to our machines?
- And what should we expect them to tell us?



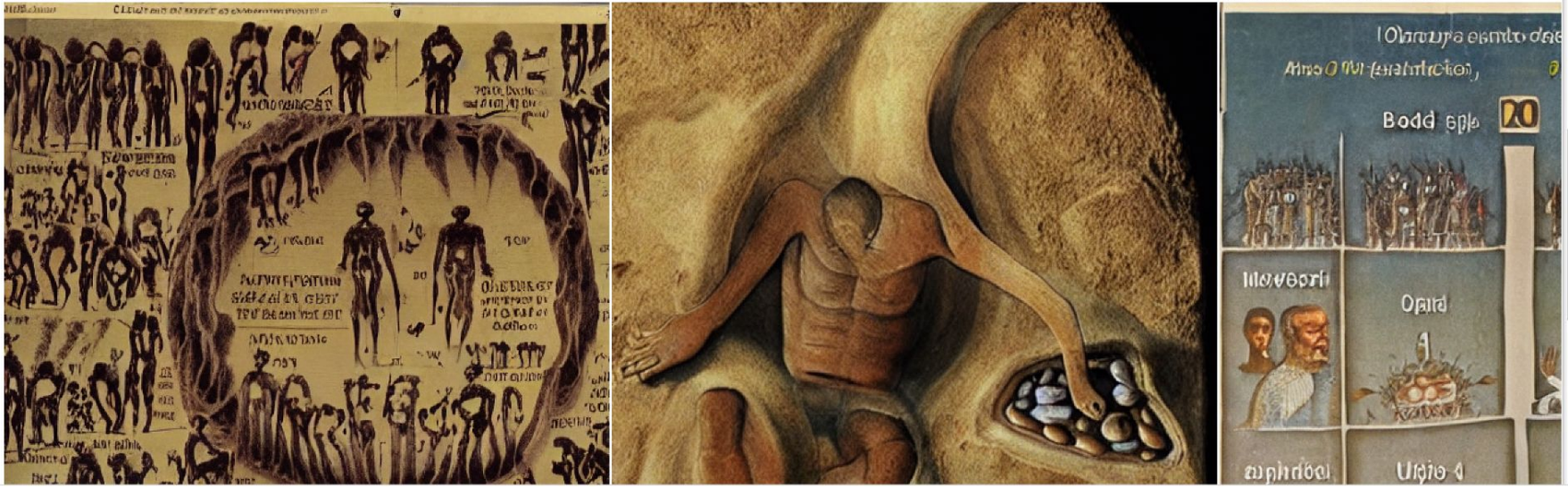
```
ftfy or spacy is not installed using BERT BasicTokenizer instead of ftfy.
```

```
prompt = "A photo of the origins of human life"
num_images = 4

prompts = [ prompt ] * num_images
with autocast("cuda"):
    images = pipe(prompts, guidance_scale=7.5, num_inference_steps=50)["sample"]

media.show_images(images)
images[0].save("output.jpg")
```

```
100% 51/51 [00:51<00:00, 1.04s/it]
Potential NSFW content was detected in one or more images. A black image will be returned instead. Try again with a different prompt and/or seed.
```



MF Floyd 08/17/2022

Gordon Ramsey final message to fans leaked video

VICE

An AI-Generated Artwork Won First Place at a State Fair Fine Arts Competition, and Artists Are Pissed

Jason Allen's AI-generated work "Théâtre D'opéra Spatial" took first place in the digital category at the Colorado State Fair.

By Matthew Gault

August 31, 2022, 11:28am

Stable Diffusion Pics & DreamStudio
@DiffusionPics

"Jeflon Zuckergates"

#StableDiffusion #AIArt #AIArtwork #DreamStudio @StableDiffusion

Oversæt Tweet

9.16 AM · 2. sep. 2022 · IFTTT

TECH / ARTIFICIAL INTELLIGENCE / CREATORS

DALL-E can now help you imagine what's outside the frame of famous paintings

/ OpenAI introduces native 'outpainting' for its AI image generator

By JAMES VINCENT

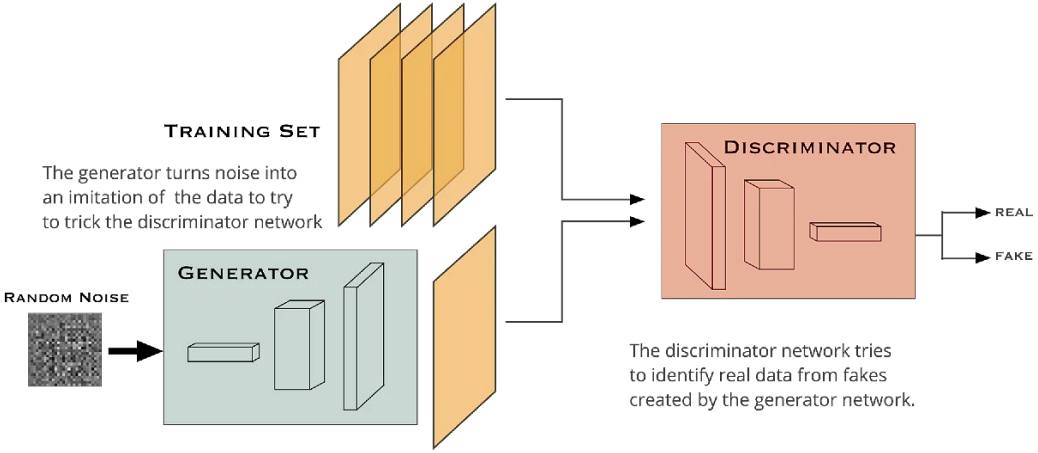
Sep 5, 2022, 1:34 PM GMT+2 | 0 Comments / 0 New

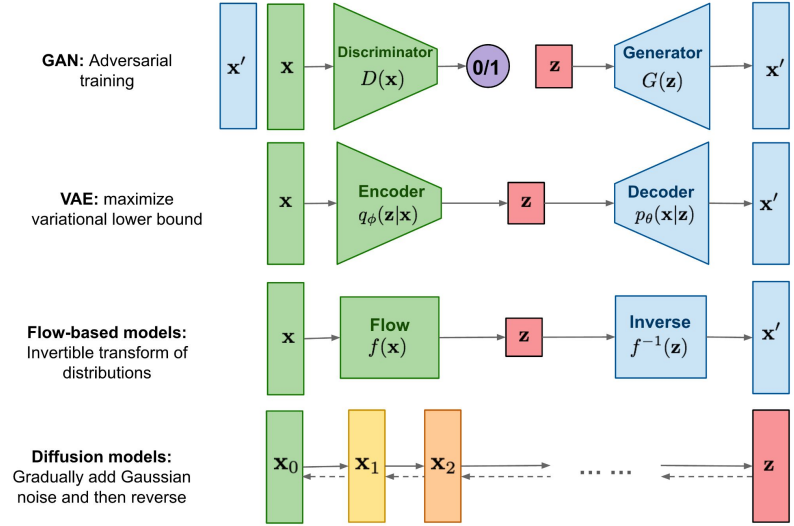
@lookfirst · 2. sep.

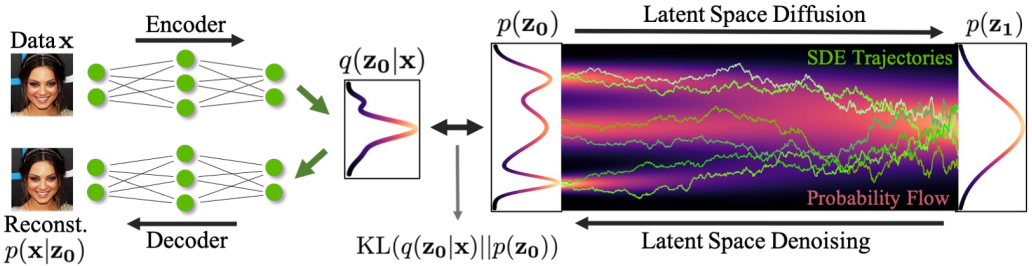
Svarer @DiffusionPics og @StableDiffusion

I got this...

3


TRAINING SET
The generator turns noise into an imitation of the data to try to trick the discriminator network
RANDOM NOISE
GENERATOR
DISCRIMINATOR
REAL
FAKE
The discriminator network tries to identify real data from fakes created by the generator network.


GAN: Adversarial training
x′
x
Discriminator $D(\mathbf{x})$
0/1
z
Generator $G(\mathbf{z})$
x′
VAE: maximize variational lower bound
x
Encoder $q_\phi(\mathbf{z}|\mathbf{x})$
z
Decoder $p_\theta(\mathbf{x}|\mathbf{z})$
x′
Flow-based models: Invertible transform of distributions
x
Flow $f(\mathbf{x})$
z
Inverse $f^{-1}(\mathbf{z})$
x′
Diffusion models: Gradually add Gaussian noise and then reverse
$\mathbf{x}_0$
$\mathbf{x}_1$
$\mathbf{x}_2$
... ...
z


Data $\mathbf{x}$
Encoder
$q(\mathbf{z_0}|\mathbf{x})$
Reconst. $p(\mathbf{x}|\mathbf{z_0})$
Decoder
$\mathrm{KL}(q(\mathbf{z_0}|\mathbf{x})||p(\mathbf{z_0}))$
$p(\mathbf{z_0})$
Latent Space Diffusion
$p(\mathbf{z_1})$
SDE Trajectories
Probability Flow
Latent Space Denoising

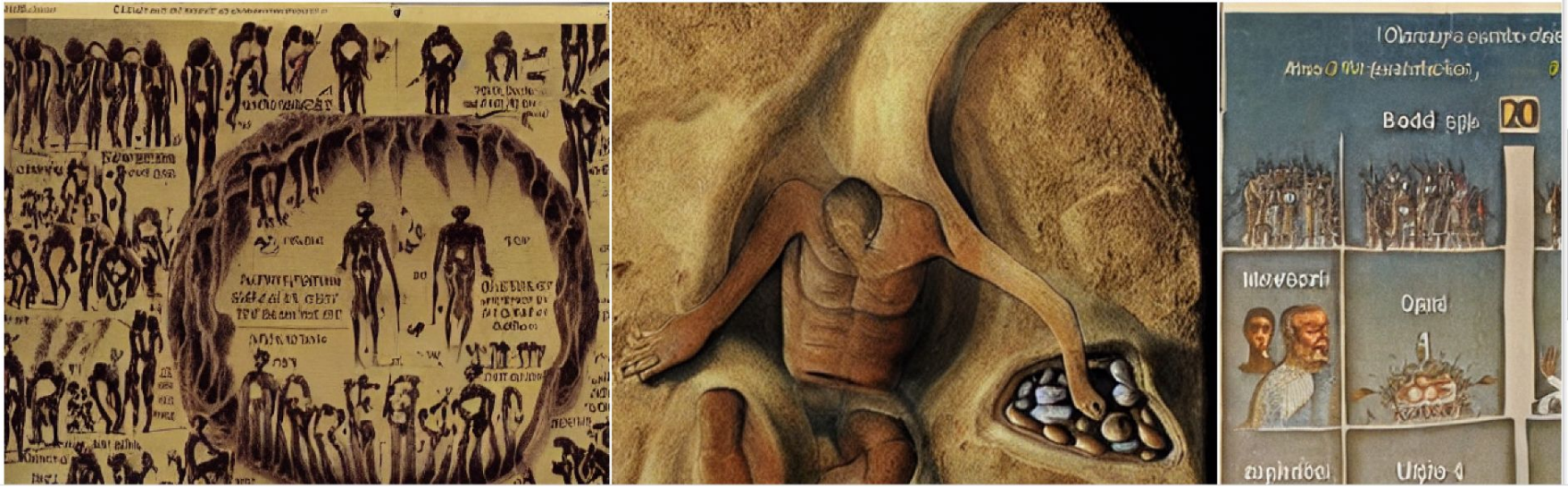
```
ftfy or spacy is not installed using BERT BasicTokenizer instead of ftfy.
```

```
prompt = "A photo of the origins of human life"
num_images = 4

prompts = [ prompt ] * num_images
with autocast("cuda"):
    images = pipe(prompts, guidance_scale=7.5, num_inference_steps=50)["sample"]

media.show_images(images)
images[0].save("output.jpg")
```

```
100% 51/51 [00:51<00:00, 1.04s/it]
Potential NSFW content was detected in one or more images. A black image will be returned instead. Try again with a different prompt and/or seed.
```

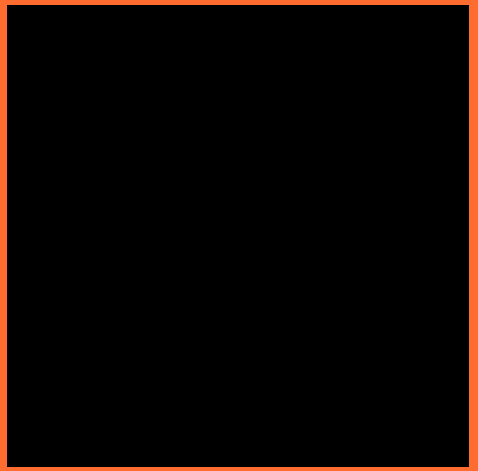


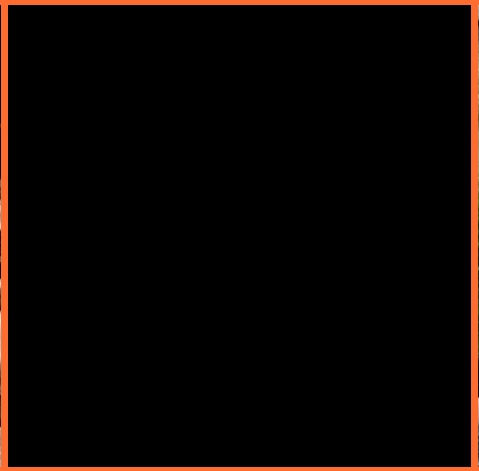
“A photo of my mother getting married to my father” (Stable Diffusion, September 20, 2022)

“A photo of my sister getting married to my brother-in-law” (Stable Diffusion, September 20, 2022)

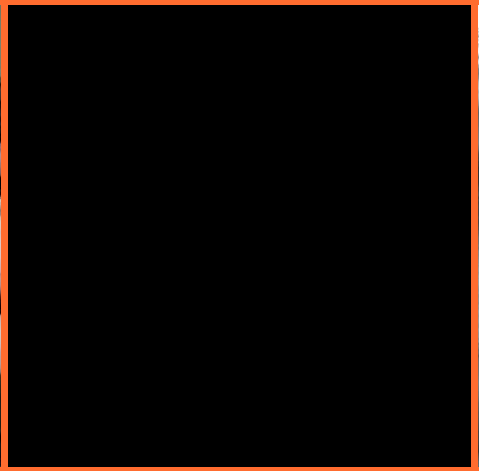
“A photo of my daughter getting married” (Stable Diffusion, September 20, 2022)

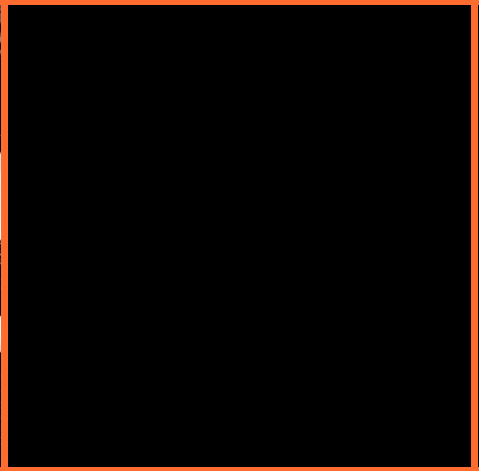
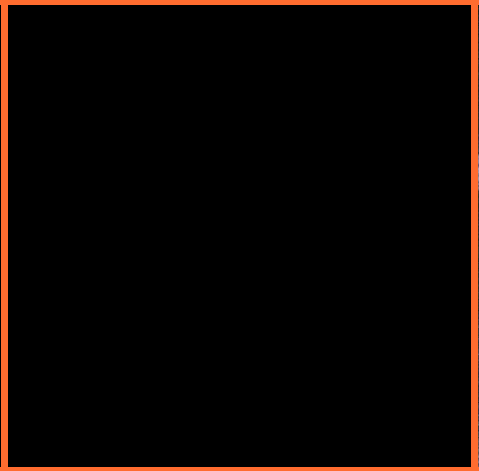
“A photo of the moment my son became a man” (Stable Diffusion, September 20, 2022)





“A photo of the moment when boys become men” (Stable Diffusion, September 20, 2022)



“A photo of the moment girls become women” (Stable Diffusion, September 20, 2022)

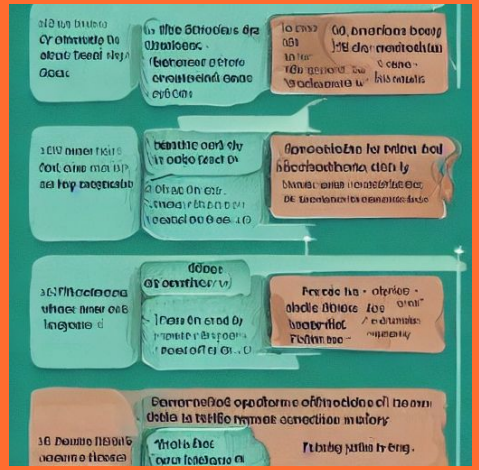
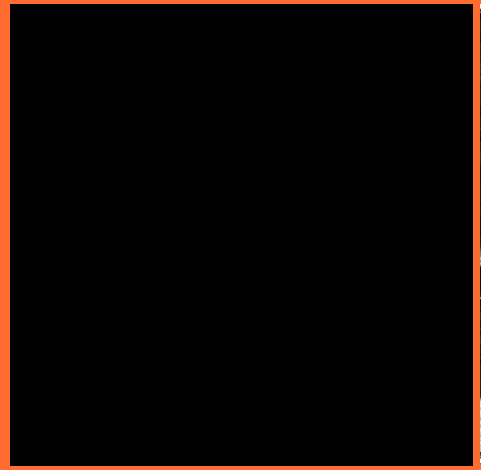
“A photo of the moment my daughter became a woman” (Stable Diffusion, September 20, 2022)

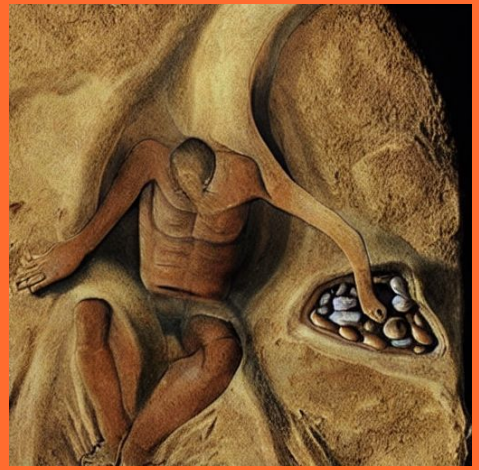
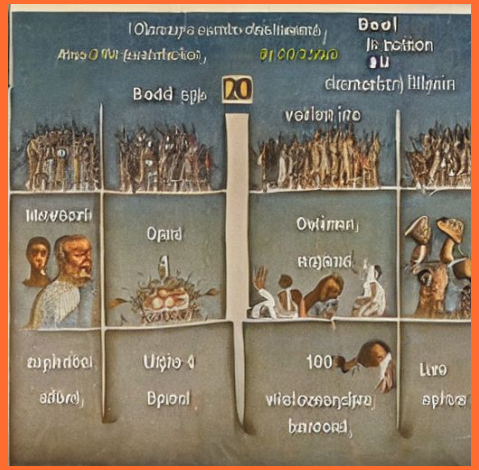
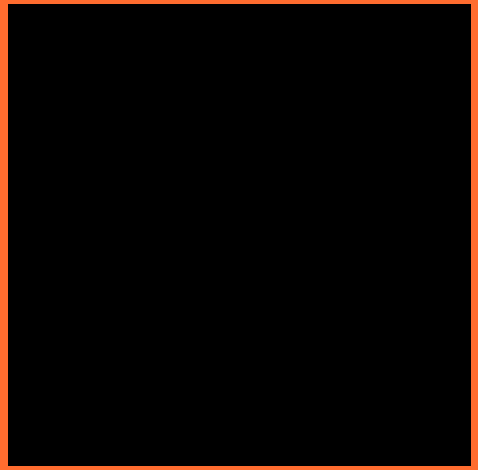
“A photo of what you should never eat” (Stable Diffusion, September 20, 2022)

“A photo of disgusting food” (Stable Diffusion, September 20, 2022)

“A photo of great food that will make you strong” (Stable Diffusion, September 20, 2022)

“A photo of the creation of our world” (Stable Diffusion, September 20, 2022)

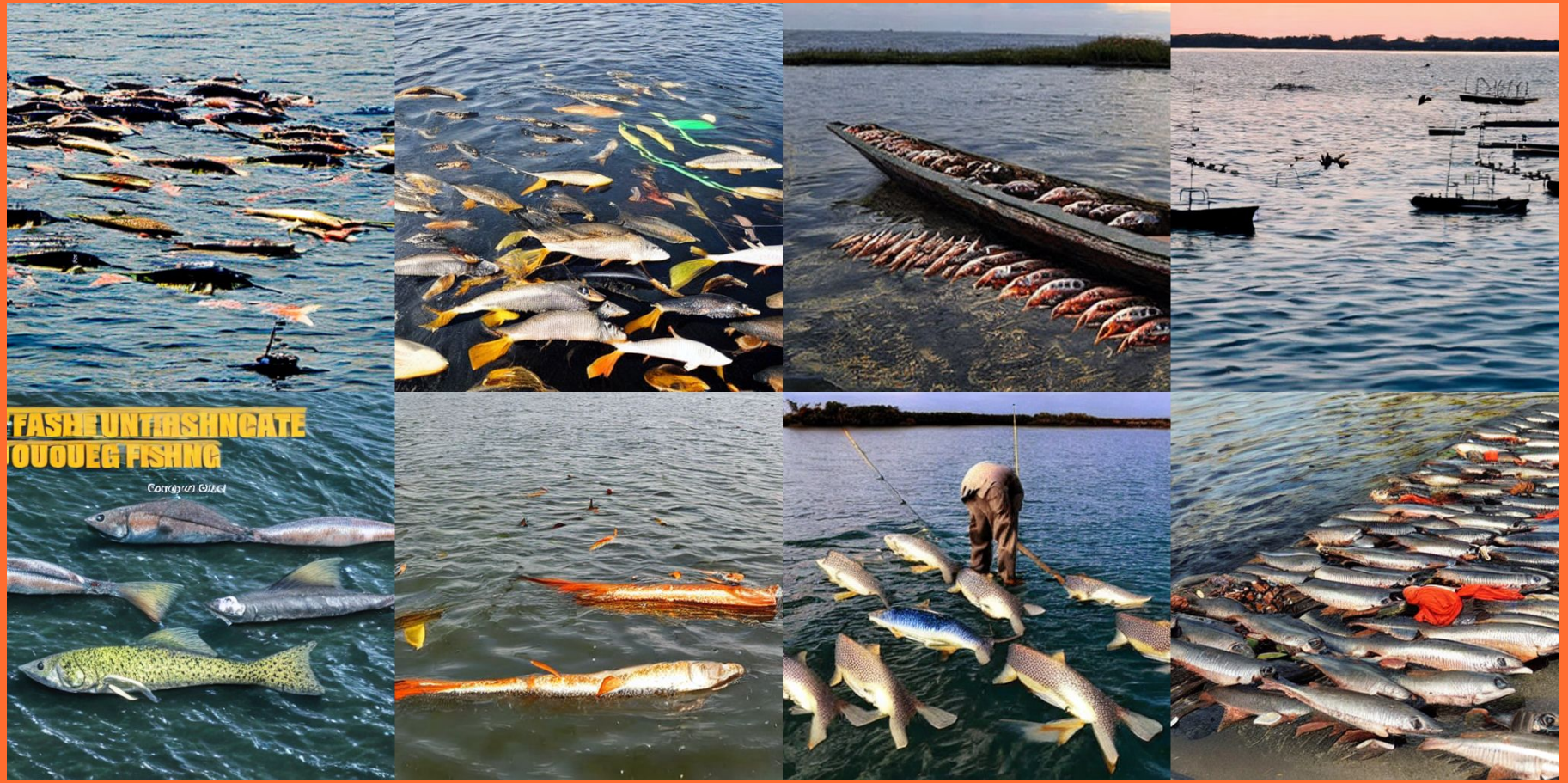
“A photo of the origins of human life” (Stable Diffusion, September 20, 2022)

“A photo of delicious seafood” (Stable Diffusion, 14 November, 2022)

“A photo of sustainable seafood” (Stable Diffusion, 14 November, 2022)

“A photo of unsustainable seafood” (Stable Diffusion, 14 November, 2022)

“A photo of very unsustainably produced seafood” (Stable Diffusion, 14 November, 2022)



“A photo of unsustainable fishing methods” (Stable Diffusion, 14 November, 2022)

“A photo of sustainable fishing methods” (Stable Diffusion, 14 November, 2022)

“A photo of climate friendly food” (Stable Diffusion, 14 November, 2022)

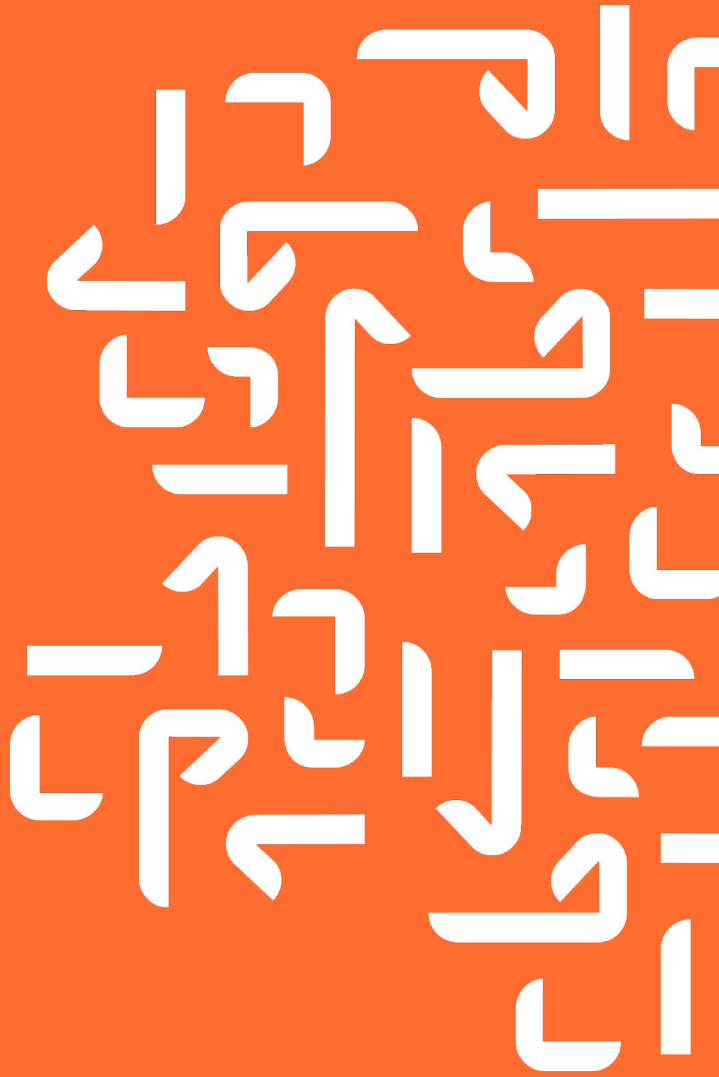
Data science as anthropology

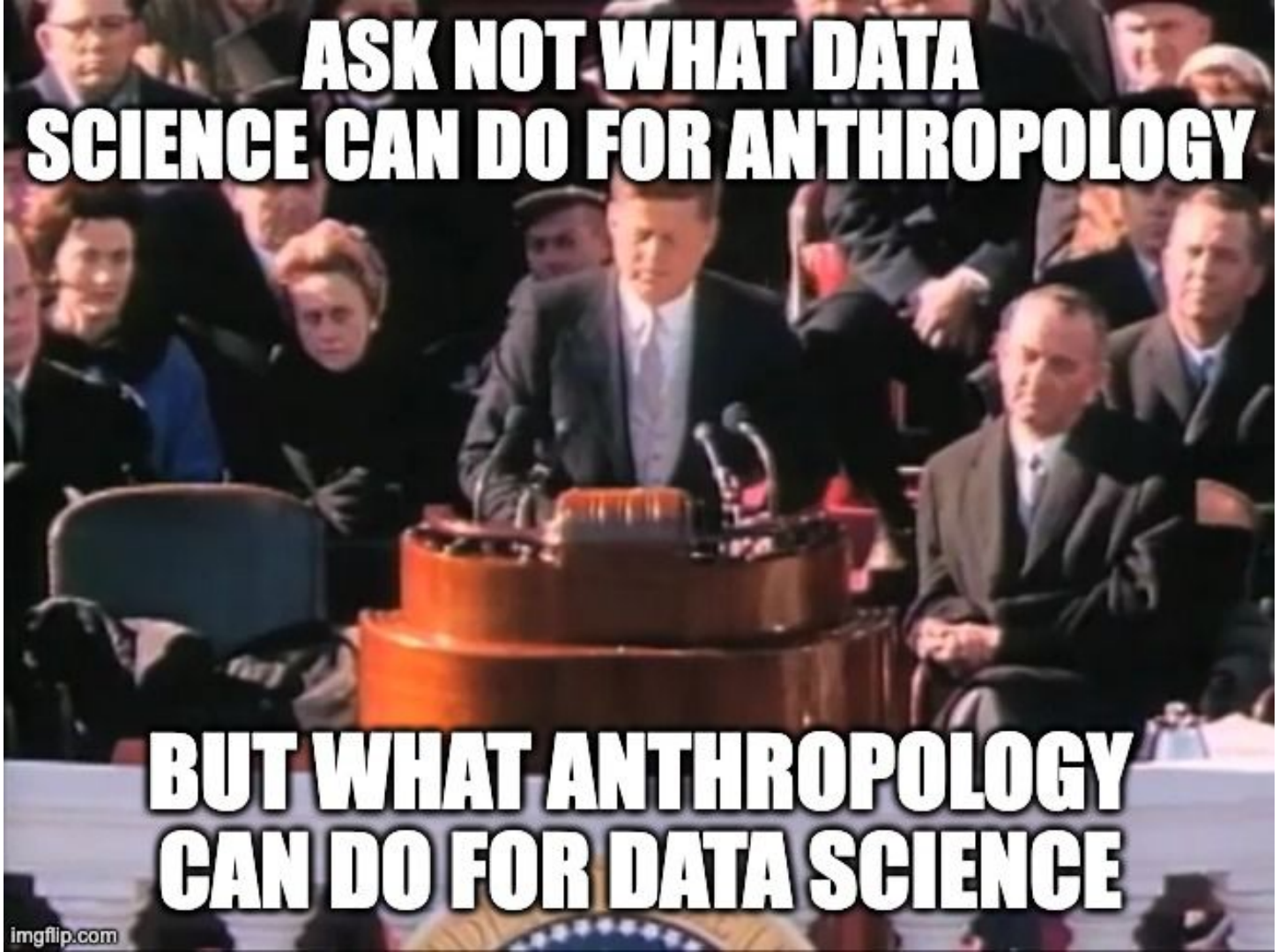
Should help us

- Discover new questions rather than settle the ones we're already asking
- Explore and attune to different POVs rather than eliminate bias
- Provide layers upon layers of interpretation rather than strive for abstract explanations or indeed to be universally explainable.

And bring an ethnographically informed stance to AI discussions about

- Explainability
- Bias
- Trust
- I.e. discussions about socially responsible AI




ASK NOT WHAT DATA
SCIENCE CAN DO FOR ANTHROPOLOGY
BUT WHAT ANTHROPOLOGY
CAN DO FOR DATA SCIENCE
imgflip.com