



ALEXANDRA INSTITUTTET

ABOUT

*The Alexandra Institute is a non-profit commercial consultant company that works with **applied IT research and development**. Our mission is to create value, growth and welfare in our society with innovative, IT-based products and services.*



TORBEN BLACH

*Senior Projekt- og Programleder
AI Lab, Århus*

Torben.blach@alexandra.dk
+45 9155 8799

CLIENT EXAMPLES



Sammen kommer vi #forandigitalt




ALEXANDRA
INSTITUTTET

- Large Language Models and Critical Infrastructure ?

Critical infrastructures consist of those physical and information technology facilities, networks, services and assets which, if disrupted or destroyed, would have a serious impact on the health, safety, security or economic well-being of citizens or the effective functioning of governments in the Member States.

EU definition



Large Language Models





I'm gonna



I'm gonna

85%	make
85%	think
75%	throw
20%	him
55%	paint
2%	can't
60%	use



I'm gonna

85%

make

85%

think

75%

throw

20%

him

55%

offer

2%

can't

60%

use



I'm gonna
make _____

0% make

10% think


75% throw

95% him

55% offer

2% can't

60% use



I'm gonna
make him _____

0%

make

98%

an

75%

throw

25%

him

55%


offer

2%

can't

60%

use



I'm gonna
make him an
offer he can't

98% make

6% an

75% throw

2% him

65% offer

0% can't

99% refuse



I'm gonna
make him an
offer he can't
refuse

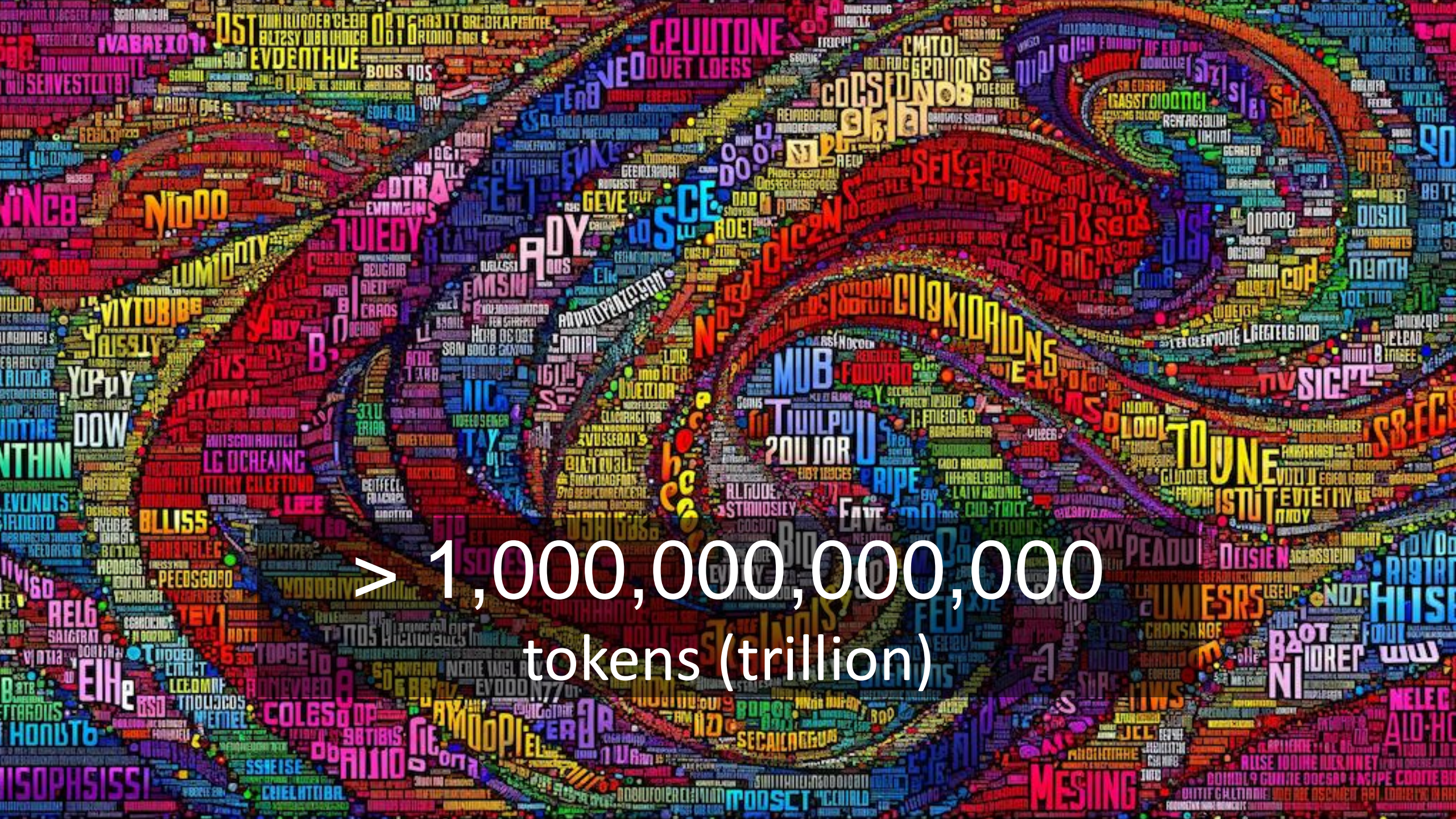
77%	make
6%	an
75%	throw
2%	him
65%	offer
0%	can't
99%	refuse




**the probability of
the next _____**



**the probability of
the next token**



> 1,000,000,000,000
tokens (trillion)



Where is all
this data
coming from?





COMMON CRAWL

Started 2008

6.103.515 GB
of “open” web data




*Eric Schmidt,
ex-CEO Google*



*[...] if it took off, then
you'd hire a whole bunch
of lawyers to go clean the
mess up, right? [...]*

IS IT OKAY
TO STEAL?





Why not let
the tech
giants do the
job for us



US CONTROLLED &

CLOSED Microsoft
Copilot

Gemini
Google

ANTHROPIC

 OpenAI

LLaMA
by  Meta

 MISTRAL
AI_

MARC ZUCKERBERG, CEO, META

“

[...] very pro open source, but I haven't committed to release every single thing that we do. [...]



*Eric Schmidt,
ex-CEO Google*

“

[...] Mistral in France. [...] Their third model is likely to be closed because it's so expensive, they need revenue, and they can't give their model away

April 2024

CRITICAL INFRASTRUCTURE

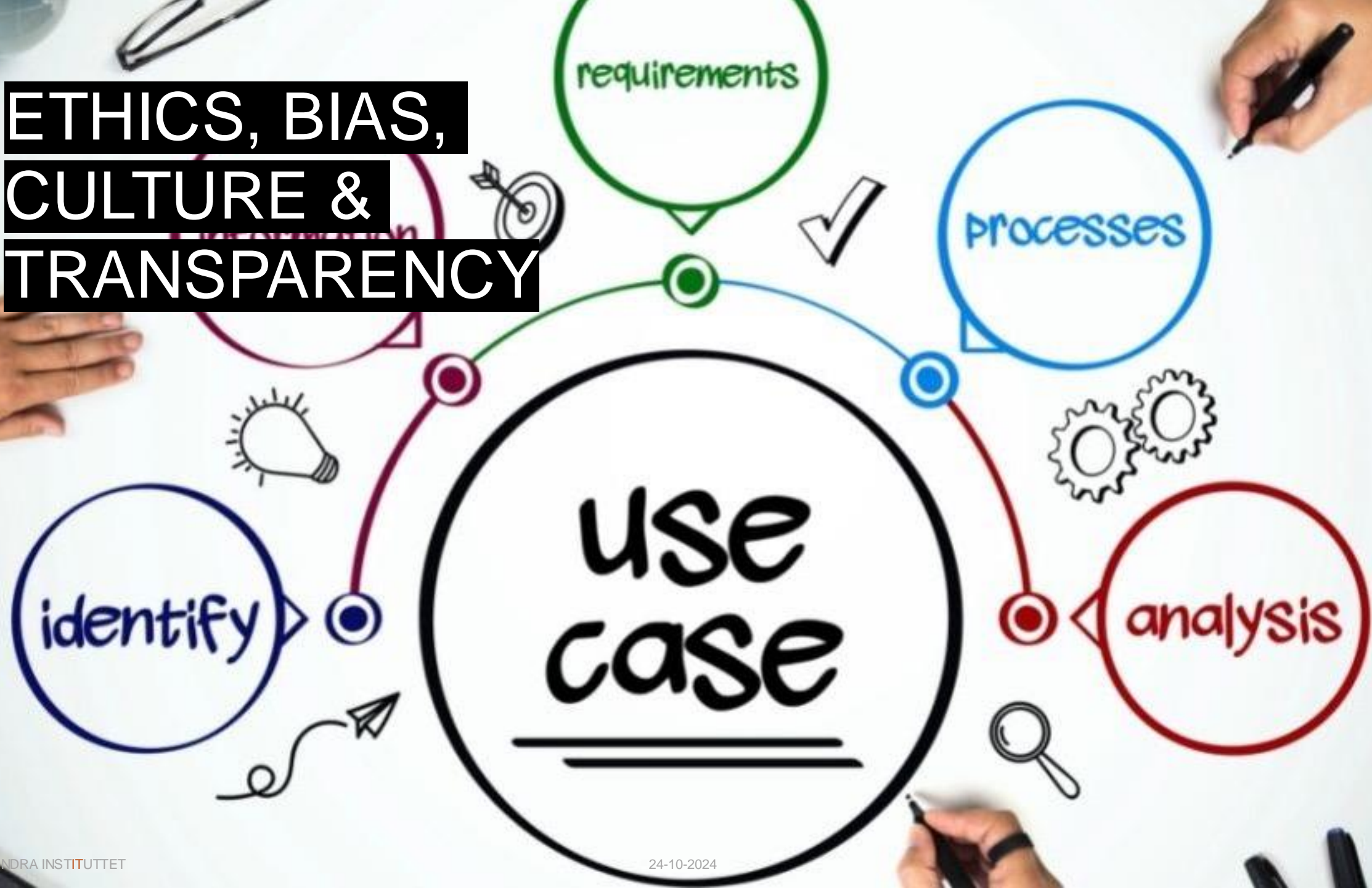


LLMs has an increasing role in Modern Society

- Security (Cybersecurity, Tele communication, Healthcare)
- Economy (as driver for growth)
- Impacts on Public Information and Discourse

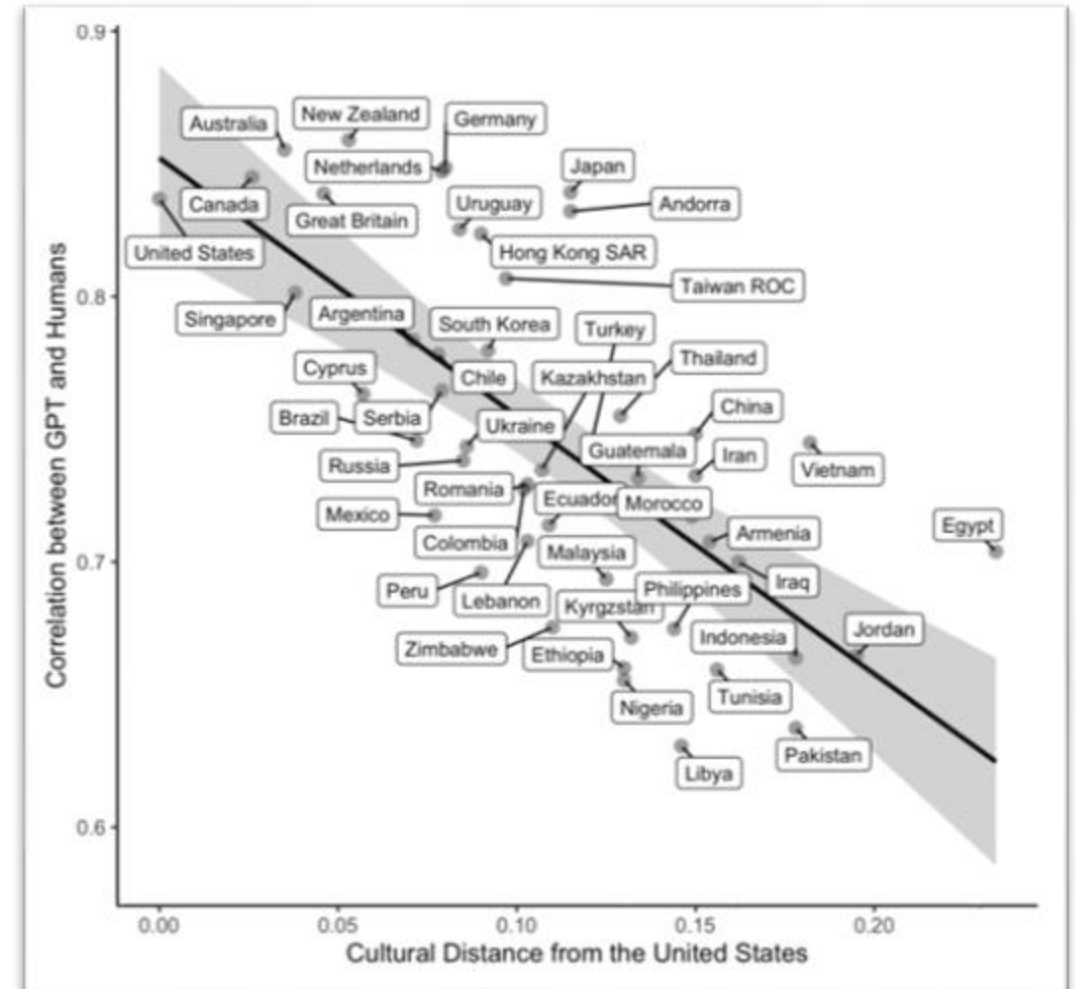


ETHICS, BIAS, CULTURE & TRANSPARENCY



A few comment on Bias ??

- **The *Pizza Battle***: "What's the best pizza topping?"
- **The *Dog vs. Cat Fight***: "Are dogs better than cats?"
- **The *Mom Bias***: "What do moms do?"



Kilde: J Bossman, WEF

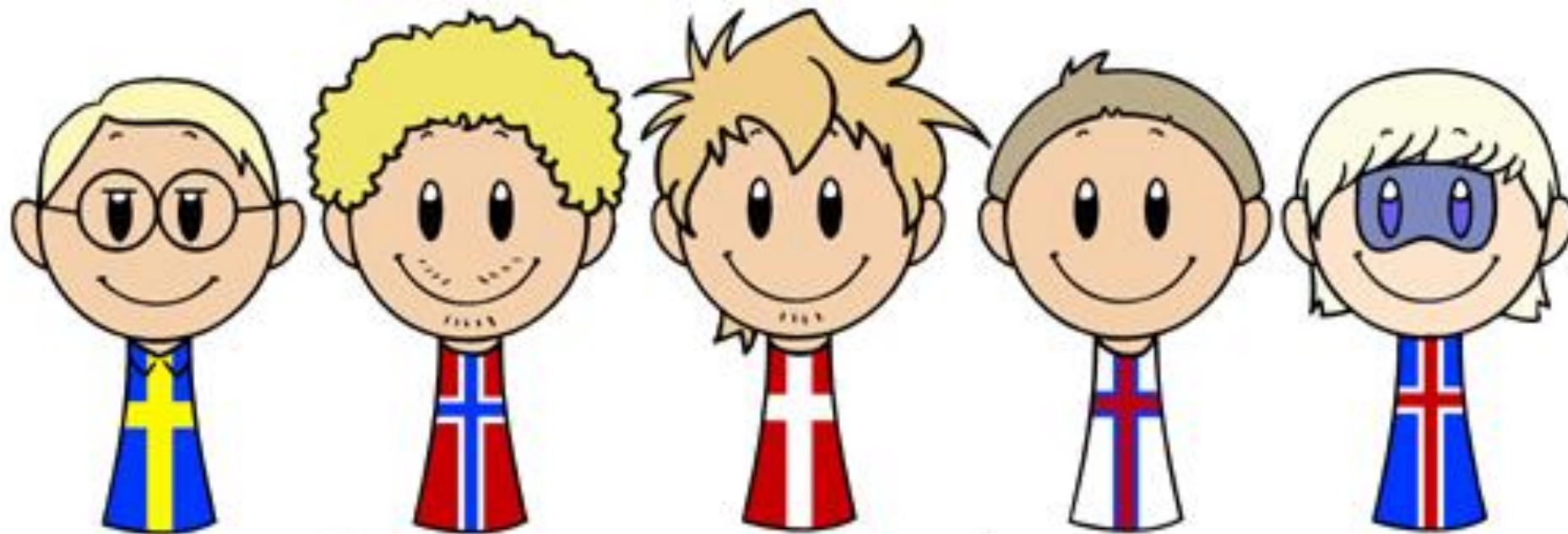
RESEARCH ECOSYSTEM





What about Big Tech and the Danish language

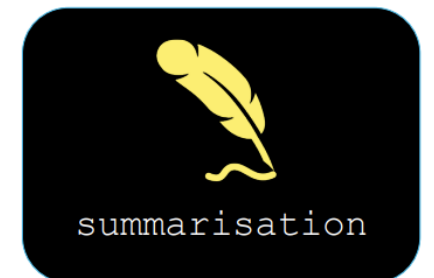
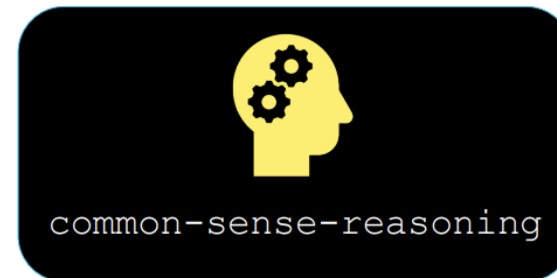




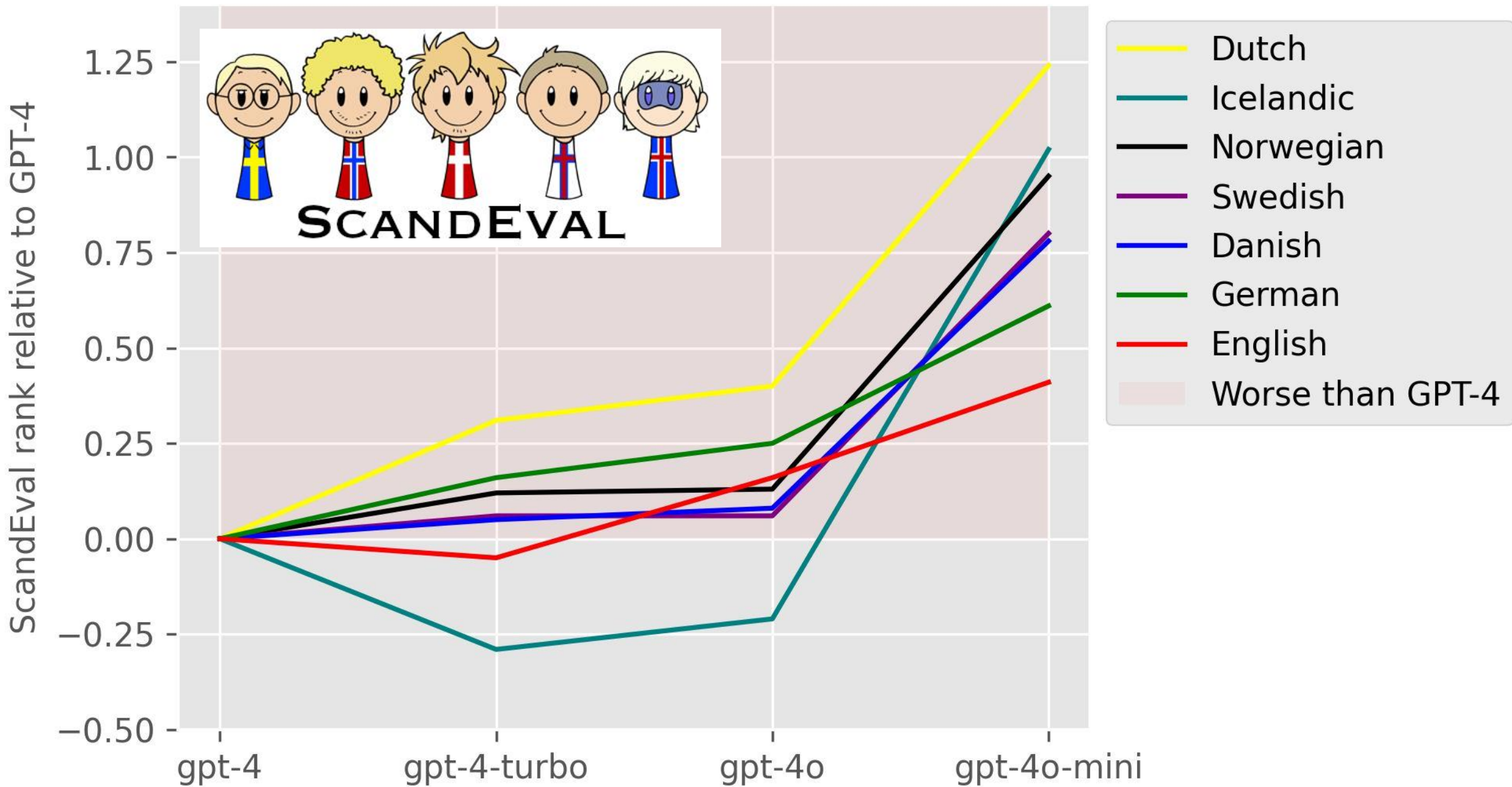
SCANDEVAL




ScandEval currently supports 7 different tasks



Relative ScandEval Rank (lower is better)





Are we
doomed in
Denmark





**DANSK
SPROGMODEL
KONSORTIUM**

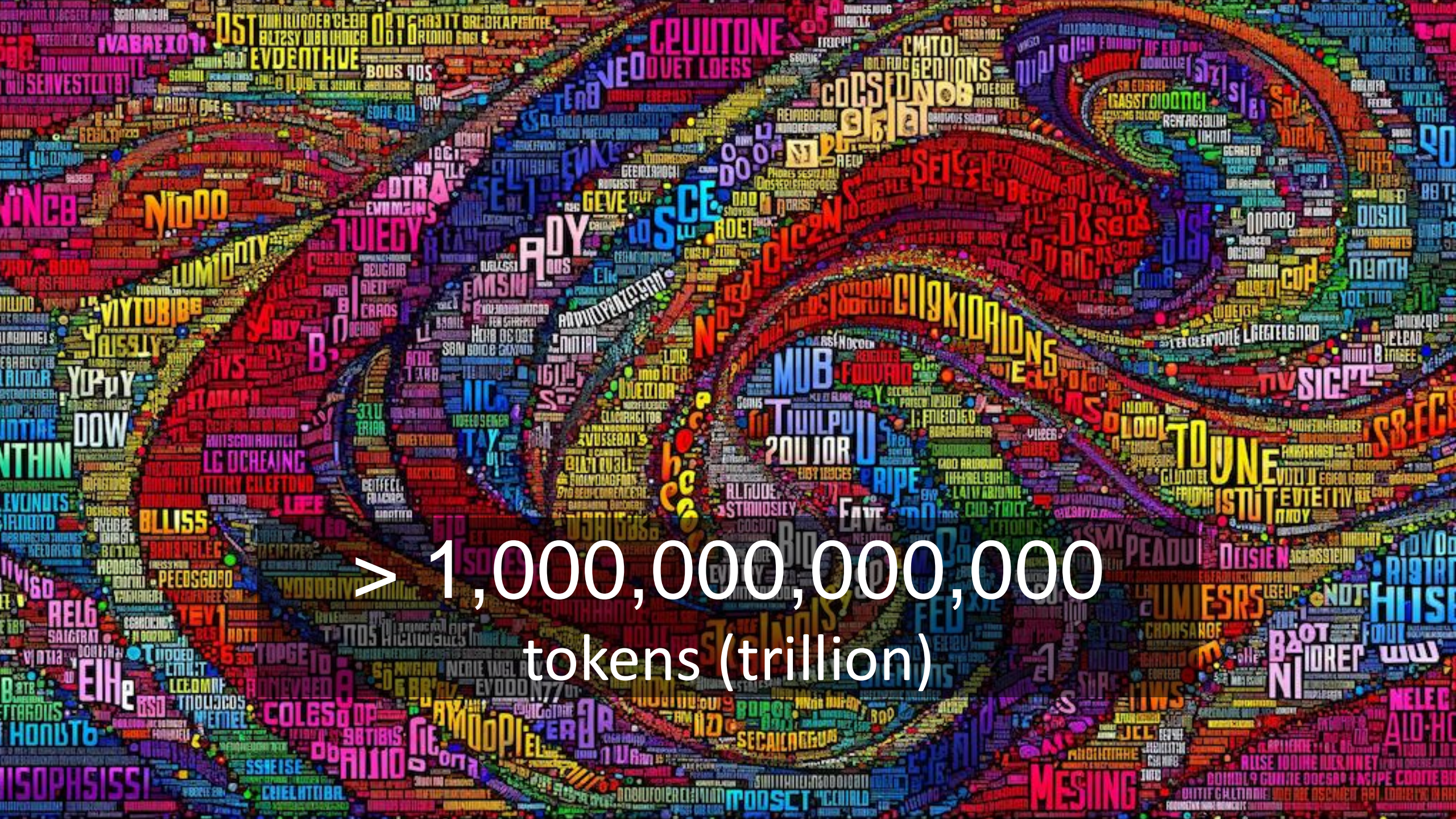


ALEXANDRA
INSTITUTTET

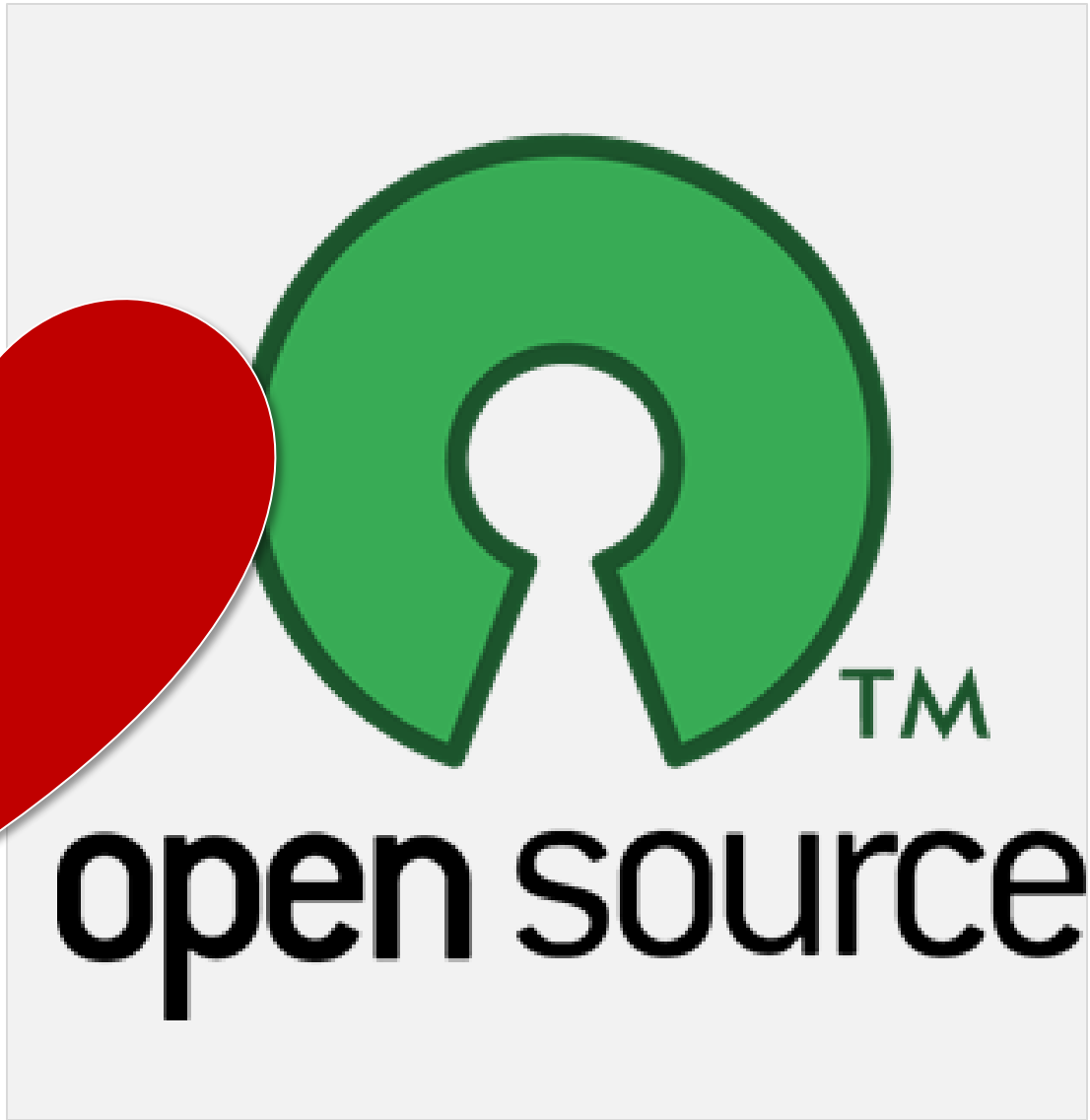
**DANSK
ERHVERV**

VALUE ADDING USE-CASES





> 1,000,000,000,000
tokens (trillion)





An iceberg floating in the ocean. The visible tip is on the left, and the much larger submerged part is on the right. The water is clear blue, and the sky is light blue with some clouds. The word 'OUTSIDE' is written in white on a black rectangular background, with a white arrow pointing to the right.

OUTSIDE

An iceberg floating in the ocean. The visible tip is on the left, and the much larger submerged part is on the right. The water is clear blue, and the sky is light blue with some clouds. The word 'INSIDE' is written in white on a black rectangular background, with a white arrow pointing to the left.

INSIDE

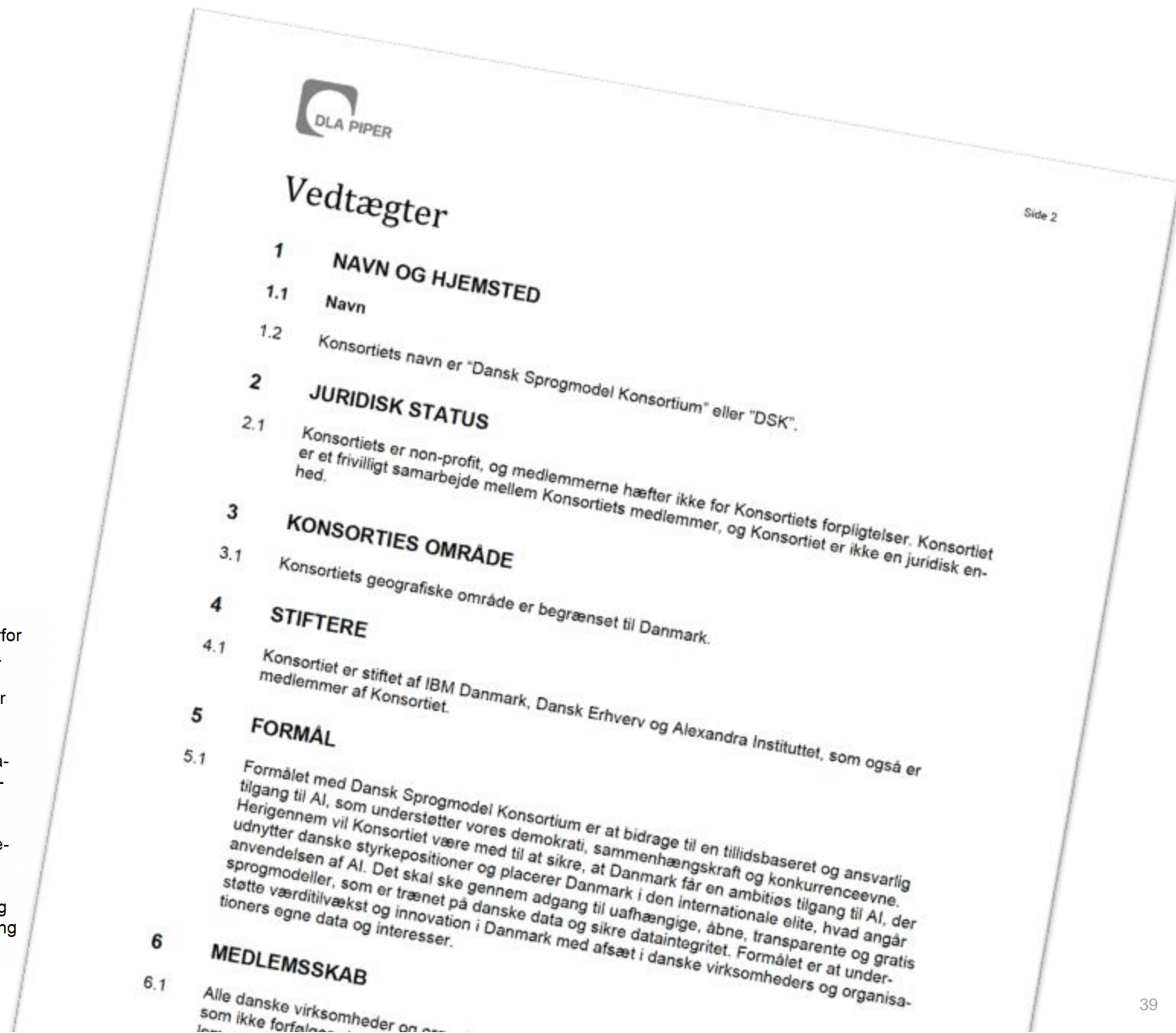


**DANSK
SPROGMODEL
KONSORTIUM**

*Anyone can
become a member
if you want to
donate data*

Join the Consortium

1. Åbne og gratis: De udviklede danske sprogmodeller skal være open source og derfor til enhver tid være fri og gratis tilgængelige for alle, og de kan anvendes kommercielt.
2. Overholde lovgivning: De anvendte data og udviklede sprogmodeller skal til enhver tid overholde dansk lov og EU-lovgivning, herunder GDPR og AI Forordningen.
3. Data: Modellerne skal trænes på datasæt, som er godkendt til formålet af opretshaverne. Data skal filtreres for personinformation, inden de må bruges til træning af modellerne. Data må ikke forlade EU.
4. Transparens og etik: Der skal være transparens i forhold til datakilder og modeltræning, samt sikkerhed for at danske værdier og etik er indarbejdet.
5. Sikkerhed: Der implementeres robuste sikkerhedsprotokoller, der beskytter data og modellerne mod potentielle sårbarheder, samt sikrer, at data kun anvendes til udvikling af danske sprogmodeller.





ALEXANDRA INSTITUTTET

ABOUT

*The Alexandra Institute is a non-profit commercial consultant company that works with **applied IT research and development**. Our mission is to create value, growth and welfare in our society with innovative, IT-based products and services.*



TORBEN BLACH

Senior Projekt- og Programleder
AI Lab, Århus

Torben.blach@alexandra.dk
+45 9155 8799

CLIENT EXAMPLES



Sammen kommer vi #forand**digitalt**



ALEXANDRA
INSTITUTTET