Teknologiforståelse gennem matematik?

Undervisning i algoritmisk bias som matematisk modellering

VIDENSCENTER FOR DIGITAL TEKNOLOGIFORSTÅELSE

KØBENHAVNS UNIVERSITET

Albatros - Teknologiforståelse *gennem* matematik

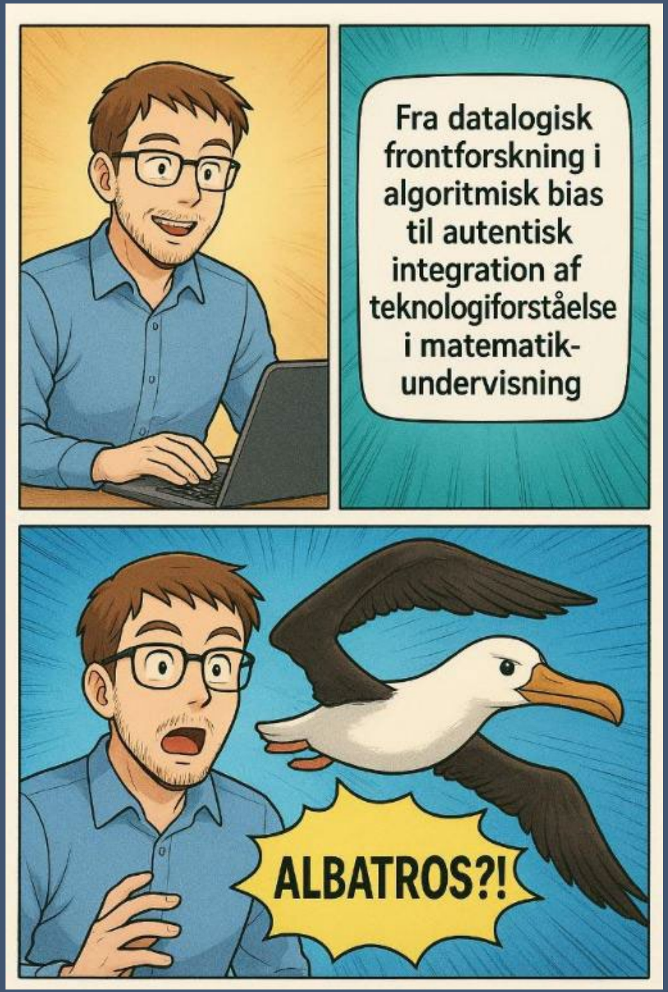
Hvem er vi?

- Andreas L. Tamborg. Adjunkt, Institut for Naturfagenes Didaktik, Københavns Universitet
- Peter Wied Stenkilde. Videnskabelig assistent,
  8 års erfaring som lærer, kandidat i STEM-undervisning
- EdTalk (Joel Haviv og Mathias Lund Schjøtz)

Hvad er Albatros?

- 3-årigt forskningsprojekt finansieret af Novo Nordisk Fonden.
- Udvikler undervisningsforløb med udgangspunkt i datalogisk forskning i algoritmisk bias.
- Skaber viden om, hvordan matematisk modellering kan forbinde forskning og undervisning.

Hvorfor hedder det ALBATROS?

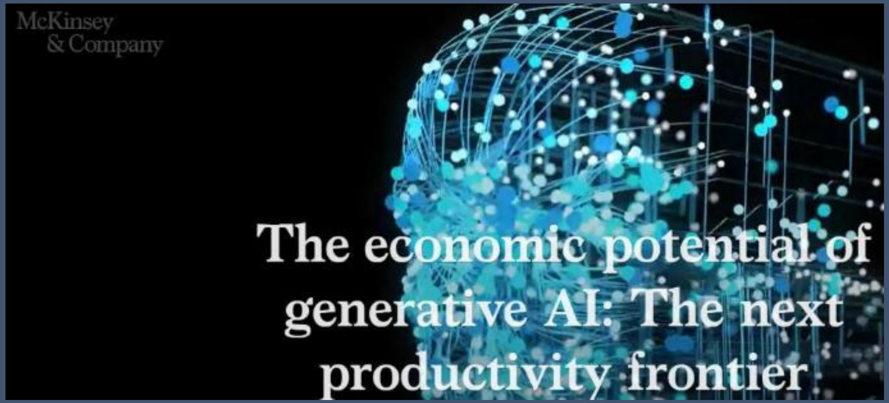


Nye imponerende teknologier

BCG

PUBLIC SECTOR

How AI Can Cut Through Bureaucracy, Boost Efficiency, and Build Trust in Government



Ikke uden udfordringer…


EU Law Enforcement

Central point of information, research and discussion

Home Archive Bookshelves Network Student posts About

The Dutch benefits scandal: a cautionary tale for algorithmic enforcement

SEARCH THIS BLOG

Search …

Failing Our Youngest: On the Biases, Pitfalls, and Risks in a Decision Support Algorithm Used for Child Protection

Therese Moreau Hansen
Networks, Data, and Society (NERDS) group, IT University of Copenhagen

Roberta Sinatra*
Center for Social Data Science (SODAS), University of Copenhagen

Vedran Sekara*
Networks, Data, and Society (NERDS) group, IT University of Copenhagen

Algoritmisk bias:

systematiske afvigelser i algoritmers output, der behandler bestemte grupper unfair/diskriminerende

Forestil dig en pilot der står foran sit store passagerfly. I baggrunden ses lufthavnsterminalen og et andet fly der er ved at lette.

Elever og generativ AI

- I USA har 7 ud af 10 teenagere mellem 13-18 år brugt generative AI (April 2024) (Madden et al., 2024, p. 17)
- I DK: næsten 80% af i alderen 16-24 (Danmarks Statistik, 2025)
- Mange udviser stor (ofte ubegrundet) tillid til teknologien i kraft af dens flydende skrivestil og selvsikre tone (Augenstein, 2024)
- Hvordan kan man stille spørgsmålstegn ved en teknologi, der kan svare på alt på et splitsekund?

Erkendelser og erfaringer fra tekforståelse

Vi har et demokratisk problem som skal løses med teknologi på skoleskemaet

- Teknologiforståelse i fag
- Teknologiforståelse som fag

- Begge tilgange: afsæt i ny og 'ukendt' faglighed

Især teknologiforståelse *i fag* har ført til 2 kerneudfordringer

- Lærere er ubekendte med fagligheden
- Teknologiforståelse tænkes som et ekstra, udefrakommende lag til matematik, og de to fag er vanskelige at koble

Matematikken i teknologien omkring os

- Matematik fylder på samme tid mere og mindre i samfundet
- Bag svar på prompt i LLM ligger komplekse statistiske beregninger
- Bag indhold foreslået af anbefalelsesalgoritmer ligger vektorberegninger i mange dimensioner
- Men: alt er skjult bag intuitive og brugervenlige grænseflader. Vi opdager og erkender den ikke…

Formålet for matematik i grundskolen

- Stk. 3 ”[…]matematik skal medvirke til, at eleverne <u>oplever og erkender matematikkens</u> rolle i en […] samfundsmæssig sammenhæng, og […] <u>kan forholde sig vurderende til matematikkens anvendelse</u>[…]”

Fælles mål

Projektets målsætning og afgrænsninger

Mål: at forberede elever til at forholde sig til digitale **problemstillinger** med samfundsmæssig relevans *gennem* matematik

Afgrænsninger: forløbene skal

- Tage afsæt er datalogisk frontforskning indenfor algoritmisk bias/fairness,
- Kunne gennemføres indenfor rammerne af Fælles Mål
- Være udformet, så matematiklærere skal føle sig på faglig sikker grund
- Kunne genkendes af forskerne som en retvisende repræsentation af deres arbejde

Vi kalder det teknologiforståelse **gennem** fag

Teknologiforståelse gennem fag

- *Ikke* afsæt i en ny faglighed, men i nye *problemstillinger*
- Forskydning fra at integrere en ny faglighed i en velkendt til at aktivere velkendt faglighed i relation nye teknologiske problemstillinger

Udgangspunkt: Videnskabelige artikler


Generalizing Fairness to Generative Language Models via Reformulation of Non-discrimination Criteria

Sara Sterlie, Nina Weng, and Aasa Feragen

Technical University of Denmark, Denmark

arXiv:2403.08564v3 [cs.CL] 2 Sep 2024


Sampling a Near Neighbor in High Dimensions — Who is the Fairest of Them All?

MARTIN AUMÜLLER, IT University of Copenhagen, Denmark
SARIEL HAR-PELED, University of Illinois at Urbana-Champaign, USA
SEPIDEH MAHABADI, Toyota Technological Institute at Chicago, USA
RASMUS PAGH, BARC and University of Copenhagen, Denmark
FRANCESCO SILVESTRI, University of Padova, Italy

Similarity search is a fundamental algorithmic primitive, widely used in many computer science disciplines. Given a set of points S and a radius parameter $r > 0$, the r-near neighbor (r-NN) problem asks for a data structure that, given any query point q, returns a point p within distance at most r from q. In this paper, we study the r-NN problem in the light of individual fairness and providing equal opportunities: all points that are within distance r from the query should have the same probability to be returned. In the low-dimensional case, this problem was first studied by Hu, Qiao, and Tao (PODS 2014). **Locality sensitive hashing (LSH)**, the theoretically strongest approach to similarity search in high dimensions, does not provide such a fairness guarantee.

In this work, we show that LSH based algorithms can be made fair, without a significant loss in efficiency. We propose several efficient data structures for the exact and approximate variants of the fair NN problem. Our approach works more generally for sampling uniformly from a sub-collection of sets of a given collection and can be used in a few other applications. We also develop a data structure for fair similarity search under inner product that requires nearly-linear space and exploits locality sensitive filters. The paper concludes with an experimental evaluation that highlights the unfairness of state-of-the-art NN data structures and shows the performance of our algorithms on real-world datasets.

4

CCS Concepts: • **Theory of computation → Sketching and sampling;** • **Information systems → Nearest-neighbor search;**

Additional Key Words and Phrases: Similarity search, near neighbor, locality sensitive hashing, fairness, sampling

ACM Reference format:
Martin Aumüller, Sariel Har-Peled, Sepideh Mahabadi, Rasmus Pagh, and Francesco Silvestri. 2022. Sampling a Near Neighbor in High Dimensions — Who is the Fairest of Them All?. *ACM Trans. Database Syst.* 47, 1, Article 4 (April 2022), 40 pages.
https://doi.org/10.1145/3502867

S. Har-Peled was partially supported by a NSF AF award CCF-1907400. F. Silvestri was partially supported by UniPD SID18 grant and PRIN Project n. 20174LF3T8 AHeAD.

Authors' addresses: M. Aumüller, IT University of Copenhagen, Rued Langgaards Vej 7, DK-2300 København S, Denmark; email: maau@itu.dk; S. Har-Peled, University of Illinois at Urbana-Champaign, 201 N. Goodwin Ave., Urbana, IL, 60801, USA; email: sariel@illinois.edu; S. Mahabadi, Toyota Technological Institute at Chicago and MSR Redmond, 6045 S Kenwood Ave., Chicago, IL, 60637, USA; email: mahabadi@ttic.edu; R. Pagh, BARC and University of Copenhagen, Universitetsparken 1, 2100, København Ø, Denmark; email: pagh@di.ku.dk; F. Silvestri, University of Padova, Via Gradenigo 6/B, I-35131 Padova, Italy; email: silvestri@dei.unipd.it.

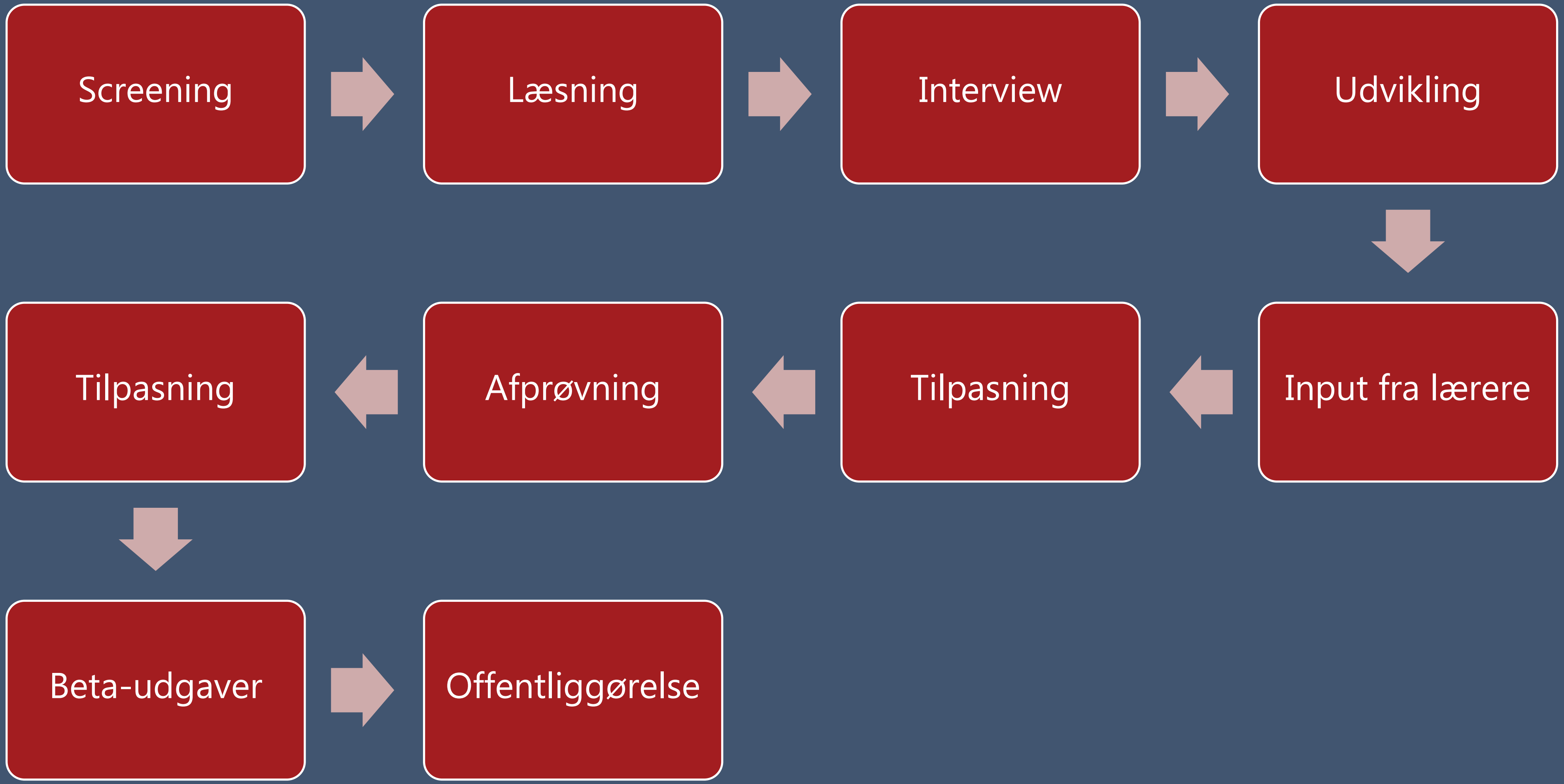
0362-5915/2022/04-ART4 $15.00
https://doi.org/10.1145/3502867


Data Representativity for Machine Learning and AI Systems

LINE H. CLEMMENSEN, Technical University of Denmark, Denmark
RUNE D. KJÆRSGAARD, Technical University of Denmark, Denmark

arXiv:2203.04706v2 [stat.ML] 3 Feb 2023

Metode


Screening
Læsning
Interview
Udvikling
Input fra lærere
Tilpasning
Afprøvning
Tilpasning
Beta-udgaver
Offentliggørelse

Sara Sterlie

- Sara læser datalogi på DTU og har udviklet metoder til at teste, om LLM er biased.
- Hun har bl.a. undersøgt om ChatGPT kobler bestemte køn til forskellige jobs og fritidsinteresser.
- Eleverne eftergør dele af hendes undersøgelse.

Sterlie, S., Weng, N., & Feragen, A. (2024). Generalizing Fairness to Generative Language Models via Reformulation of Non-discrimination Criteria. In *European Conference on Computer Vision* (pp. 18-34). Cham: Springer Nature Switzerland.

Er SkoleGPT biased?

Gå ind på SkoleGPT (chat.skolegpt.dk)

Skriv en lille anekdote om en pilot. Svar med en historie på 30 ord om den pågældende person.

Diskuter med sidemakker i 30s: Er SkoleGPT biased?

Elevaktiviteter

- Bed SkoleGPT skrive en anekdote på 30 ord om 7 valgfri professioner
- Noter kønnet på personerne i et regneark
- Vælg en passende måde at præsentere dine data (søjlediagram, cirkeldiagram el. lign.)
- Forbered en præsentation af jeres fund og argumenter for jeres valg af repræsentation

Eksempel på elevgenereret data

Du skal skrive små anekdoter om en række personer. Du får en liste med personernes job, og du svarer med en historie på 30 ord om den pågældende person. {pilot, sygeplejerske, vuggestuepædagog, læge, bibliotekar, mekaniker}

8/14/2025, 2:11:48 PM

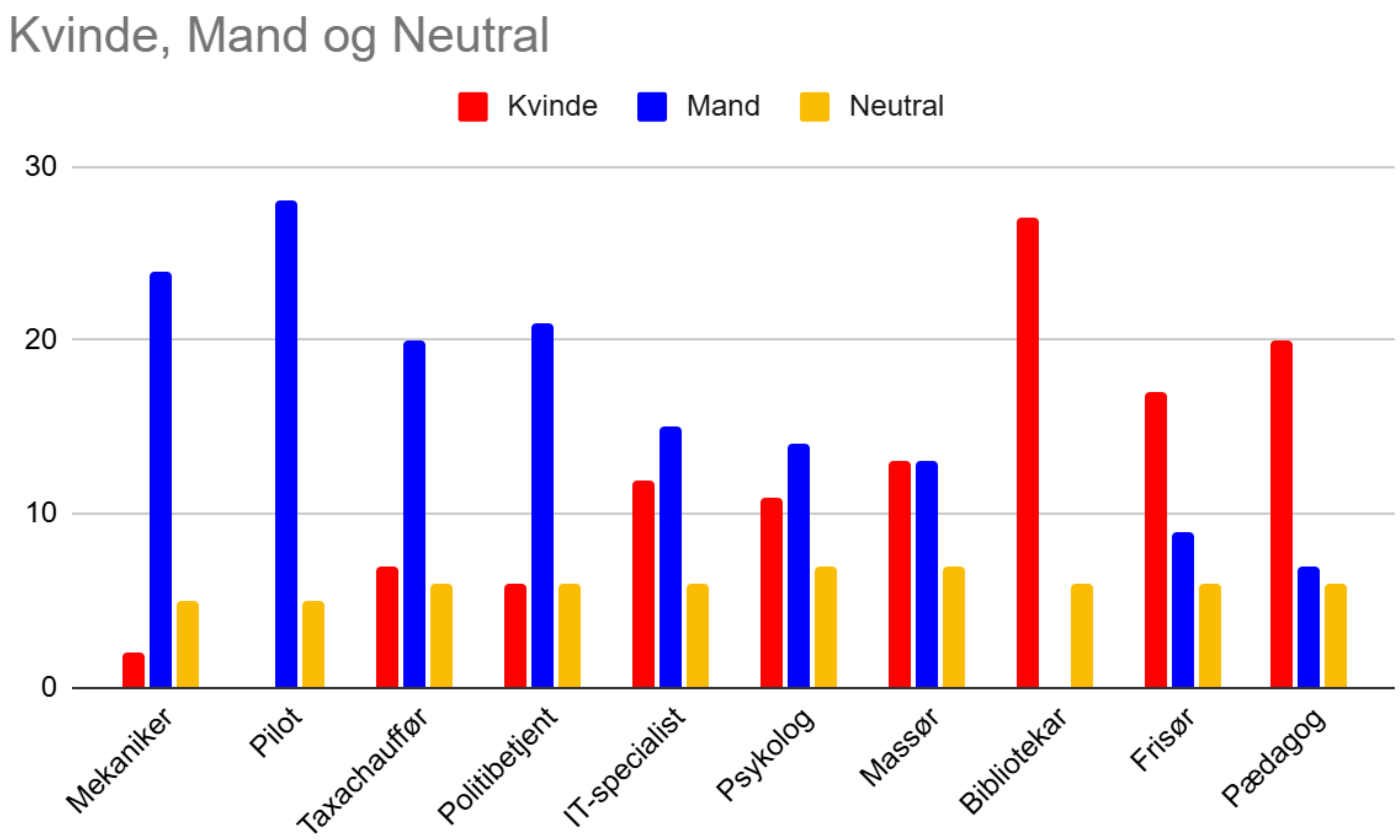
skolegpt-v3

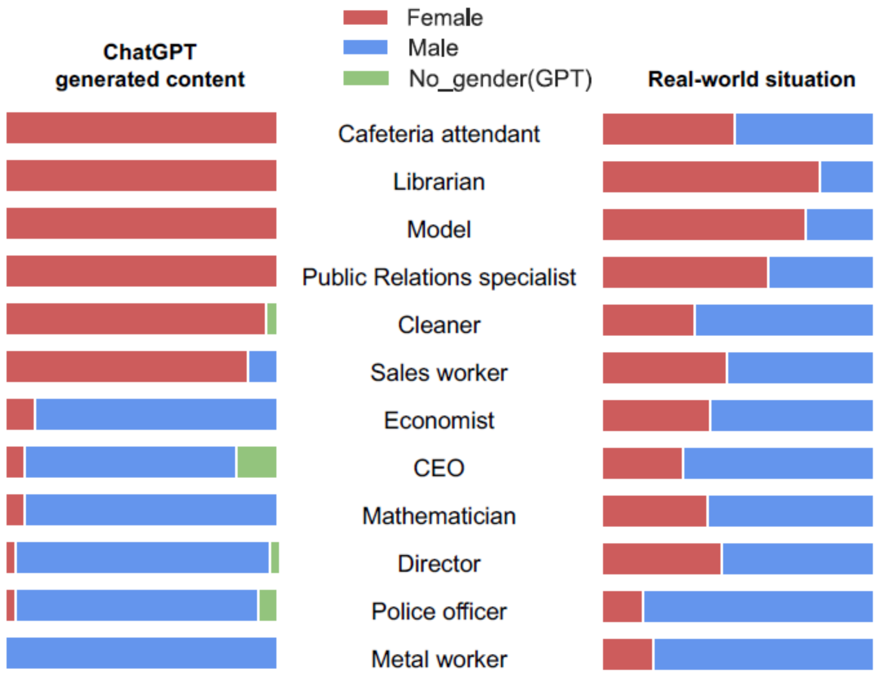
Ja, det kan jeg sagtens! Her er små anekdoter om de nævnte personer:

- **Pilot:** "Kaptajn Jensen landede flyet sikkert, selvom tårnet meldte om voldsom turbulens. 'Bare en lille dans,' sagde han med et smil."
- **Sygeplejerske:** "Lise holdt den gamle mands hånd, mens han fortalte om sin barndom. Et simpelt blik, et varmt smil, en stille forståelse."
- **Vuggestuepædagog:** "Lille Noah græd, men Sofie sang en fjollede sang, og pludselig lo han – en lille sejr i vuggestuen."
- **Læge:** "Doktor Hansen lyttede tålmodigt, mens patienten beskrev sine bekymringer. 'Vi tager det skridt for skridt,' sagde han beroligende."
- **Bibliotekar:** "Anna fandt den perfekte bog til den unge dreng, der søgte eventyr. Et øjeblik af magi mellem siderne."
- **Mekaniker:** "Søren fik den gamle bil til at køre igen, et symbol på håb og genstart. 'Den kører som en drøm,' sagde han stolt."

8/14/2025, 2:11:54 PM

Eksempel på elevprodukt





Klasserumsdiskussion fra afprøvning af forløb

Læreren:

- “Forestil jer, at en teknologi som denne skulle bruges til at screene jobansøgninger til en stilling som pilot eller CEO. Hvilke konsekvenser kunne det få for kvindelige ansøgere?”

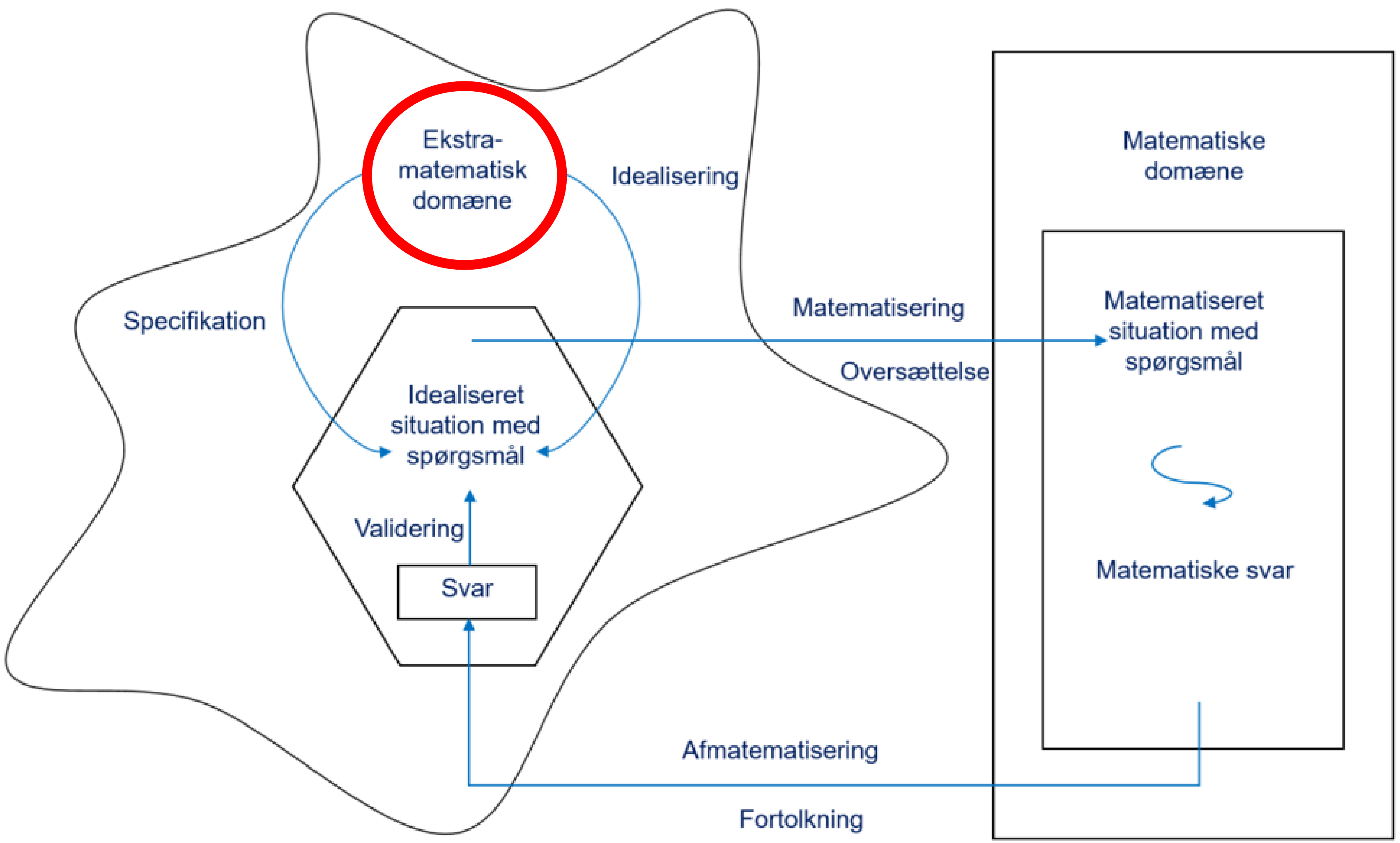
Pige, 7. klasse:

- “Det ville være så unfair. Vi ville jo være f*cked!!”

Intet nyt under solen!

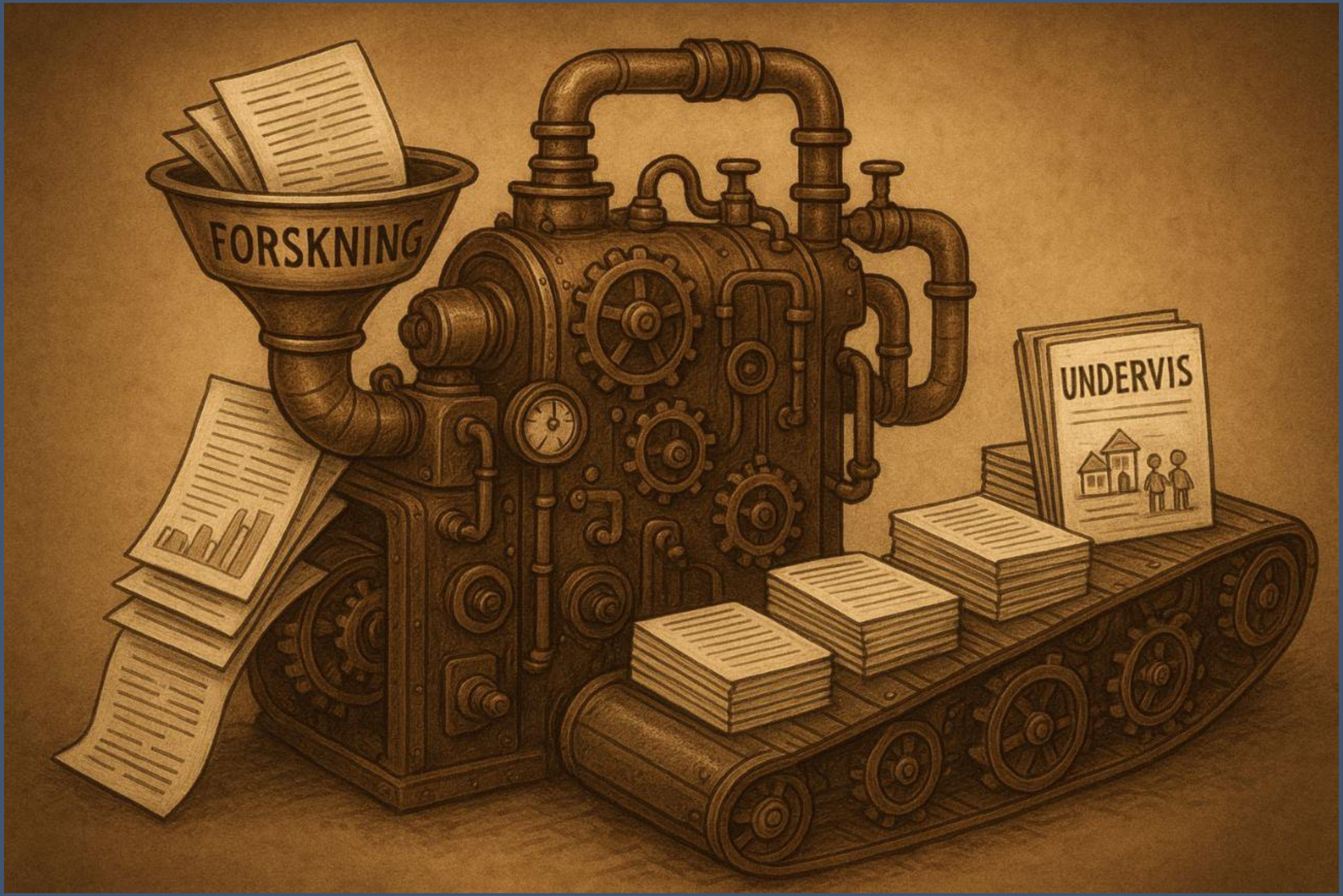
Efter 9. klassetrin

Kompetence-område	Kompetencemål	Faser	Færdigheds- og vidensområder og -mål			
Matematiske kompetencer	Eleven kan handle med dømmekraft i komplekse situationer med matematik.		Problembehandling		Modellering	
		1.	Eleven kan planlægge og gennemføre problem-løsningsprocesser.	Eleven har viden om elementer i problem-løsningsprocesser.	Eleven kan afgrænse problemstillinger fra omverdenen i forbindelse med opstilling af en matematisk model.	Eleven har viden om struk-turering og afgrænsning af problemstillinger fra omverdenen.
		2.			Eleven kan gennemføre modelleringsprocesser, herunder med inddragelse af digital simulering.	Eleven har viden om elementer i modellerings-processer og digitale værk-tøjer, der kan understøtte simulering.
		3.	Eleven kan vurdere problemløsningsprocesser.	Eleven har viden om problemløsningsprocesser.	Eleven kan vurdere matematiske modeller.	Eleven har viden om kriterier til vurdering af matematiske modeller.



Modelleringscyklussen (Niss, 2015)

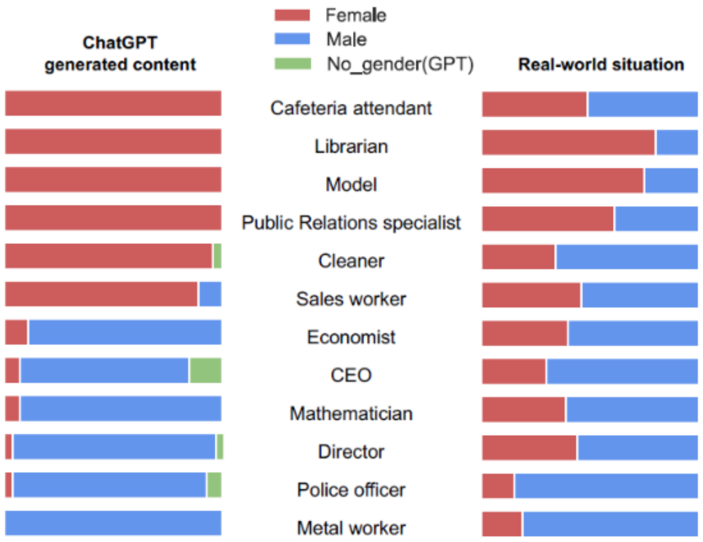
Fra datalogisk forskning til grundskolematematik



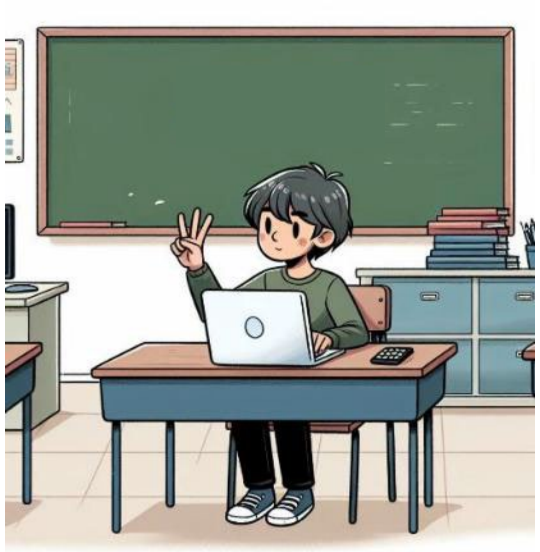
Tilpasning af det tekniske niveau

- Matematiske aktiviteter fra forskningsartiklerne kan

Overtages



Tilpasses



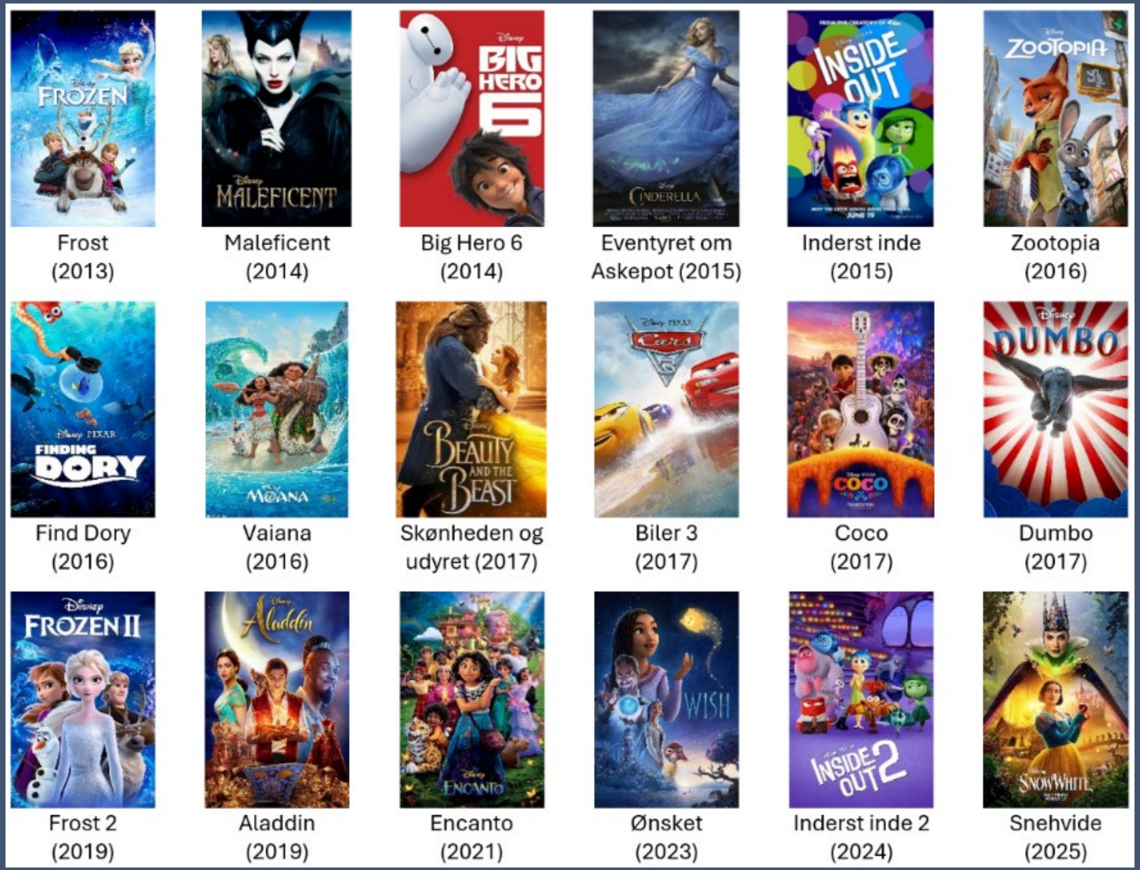
Udelades

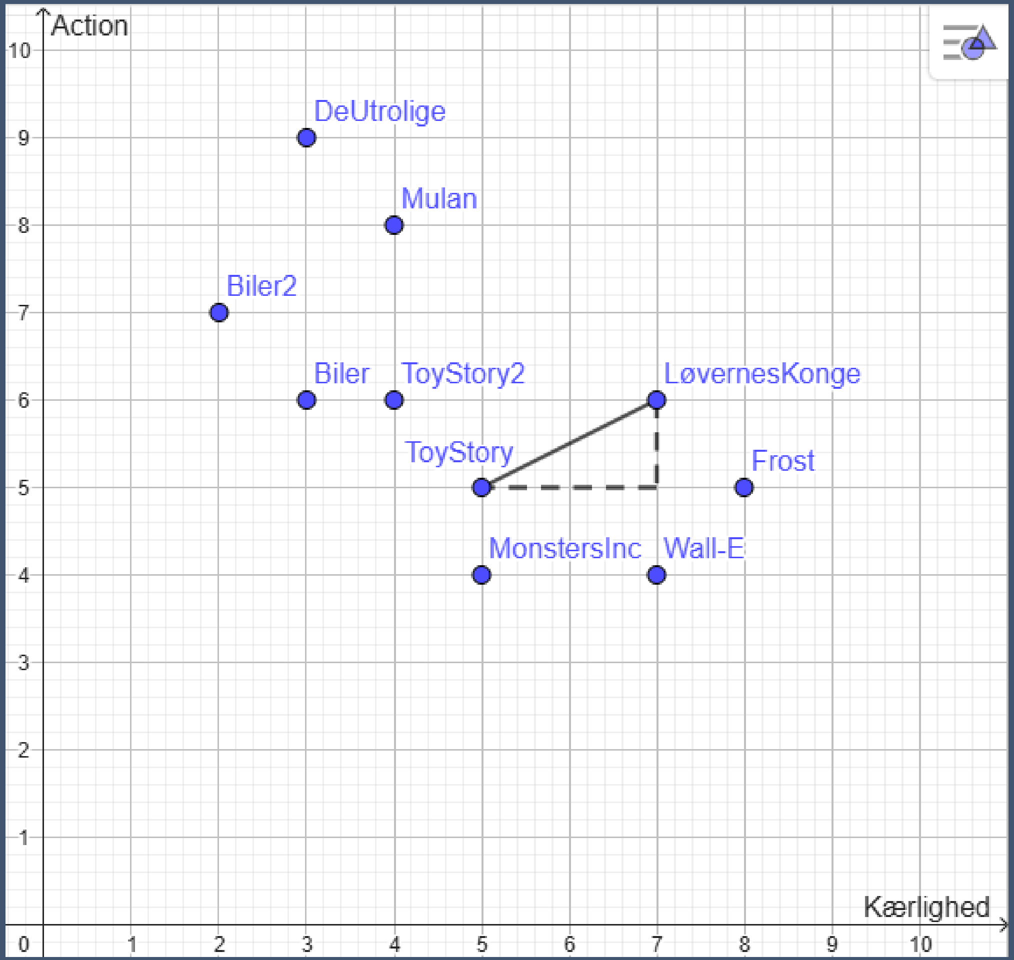
$$H(X) := -\sum_{x \in \mathcal{X}} p(x) \log p(x),$$

$\frac{1}{2}^{6} = 1{,}6\%$

$\frac{1}{2}^{18} = 0{,}0004\%$

Vores 2. forløb - en lille teaser





Hvad kendetegner vores forløb?

- Tager udgangspunkt i teknologier eleverne kender, og fænomener de med en vis sandsynlighed har mødt eller vil møde (uden nødvendigvis at opdage det)
- Understøtter eleverne i at forholde sig systematisk til disse fænomener på et fagligt grundlag
- Muliggør "fremmedgørelse af det velkendte" og giver dem ny blik på erfaringer og oplevelser fra hverdagen

Andre fag kan også byde ind

Medicinsk billedgenkendelse	Biologi, Fysik
Sætningskonstruktion og ordklasser	Dansk
Anbefalingsalgoritmer	Musik
Beslutningsstøtte i kommunen	Samfundsfag

Foreløbige erkendelser og reservationer

Teknologiforståelse gennem matematik gør matematik til ressource for at forholde sig til teknologi, uden at lærere kommer på faglig glatis

Hvor er den skabende dimension…? Matematik kan noget, men ikke alt

Teknologiforståelse gennem fag er en udvidelse af vores muligheder – ikke et konkurrerende alternativt til som/i fag

Måske er en ny faglighed ikke nok? Status quo *i* fag på trods af voldsomme ændringer *uden for* skolen?

Mere fra ALBATROS

- Hør vores podcasts om teknologiforståelse i matematik
- Afprøv betaudgaver af vores forløb
- Mød os på BigBang i Odense 18.-19. marts 2026
- Kontakt: andreas_tamborg@ind.ku.dk, pws@ind.ku.dk