Gefion AI supercomputer

Morten Bache, Scientific Director
(on behalf of the Gefion project team)
Novo Nordisk Foundation

Gefion. Norse goddess associated with virginity, fertility, agriculture, prosperity, and happiness. The name might literally mean “the giving one”, and she is closely associated with the goddess Freya.

NVIDIA®

DCAI DANISH CENTRE FOR AI INNOVATION

novo nordisk **foundation**

Gefion and Danish Center for AI Innovation

- A national AI partnership with NVIDIA


Berlingske | 2. SEKTION | 20.03.2024

Business

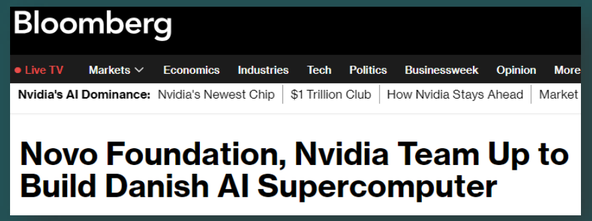
RENTEPOLITIK
Ulrik Bie: Japan dropper ære med negative renter som sidste land
SIDE 10

DIGITAL
Nu ringer det ud til sidste omgang: **TDC slukker for fastnettet** efter 142 år
SIDE 14-15

TALENT 100
Microsoft-direktør drikker **kaffe med de unge** og spørger, hvad hun kan gøre
SIDE 12-13

Ny gigantcomputer skal give dansk sundhedsforskning superkræfter

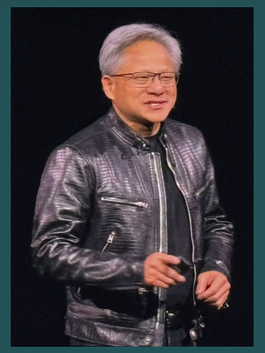
Et nyt teknologisk gennembrud for Danmark blev i går præsenteret på Statens Museum for Kunst med erhvervsminister Morten Bødskov (S) og topchef for Novo Nordisk Fonden Mads Krogsgaard Thomsen i front. Kongen deltog også for at fortælle om den nye supercomputer, der kan forandre verden.
SIDE 4-8


Bloomberg

Live TV | Markets | Economics | Industries | Tech | Politics | Businessweek | Opinion | More

Nvidia's AI Dominance: Nvidia's Newest Chip | $1 Trillion Club | How Nvidia Stays Ahead | Market

Novo Foundation, Nvidia Team Up to Build Danish AI Supercomputer



Danish announcement

...with King Frederik and Danish ministers at SMK

International announcement

...at the NVIDIA GTC conference in Silicon Valley, USA

Gefion and Danish Center for AI Innovation

- A national AI partnership with NVIDIA

Research

Research collaboration

- Pharmaceuticals and biotechnology
- Human centric AI healthcare solutions
- Acceleration of green transition
- Fault tolerant quantum computing

Infrastructure

Establishment of a national AI Supercomputer in Denmark

- Provides researchers with access to state-of-the-art AI computing services
- NVIDIA DGX SuperPOD reference architecture
- 1528 H100 GPUs and 7 PB storage
- Access to NVIDIA software platforms, training and expertise
- Secure cloud services to be established with DK partners

DCAI
DANISH CENTRE FOR AI INNOVATION

EIFO

EIFO
DKK 100 million
(EUR 13 million)

novo nordisk foundation
Benefitting people and society

NNF initial commitment
DKK 600 million
(EUR 81 million)

NNF Data Science Initiative 2020 – today

Purpose: to capacity build within data science and artificial intelligence and to support education and training of the next generation of data scientists

Open competition grants

Collaborative Research

Grants for collaborative projects involving data science within the Foundation's focus areas

Investigator Grants

Funding for excellent independent data science group leaders at different career stages

Research Infrastructure

Funding for shared super computers, hardware, GPUs, equipment, and "data as infrastructure"

Standalone initiatives

Data Science Academy

Educational courses, Symposia and networking, PhD & post-doc grants, Visiting Professorships, Vocational Training

AI Pioneer Center

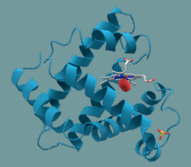
State-of-the-art AI research

Compute infrastructure?

Not funded under current initiative

Motivation: why is NNF investing in an AI Supercomputer?

AI is a key enabler of NNF's strategy...

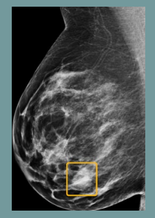


Medical research

AI shows promise for breast cancer screening

Etta D. Pisano

Could artificial intelligence improve the accuracy of screening for breast cancer? A comparison of the diagnostic performance of expert physicians and computers suggests so, but the clinical implications are as yet uncertain. See p.89



Review

Application of AI techniques and robotics in agriculture: A review

NEWS | 20 February 2020

Powerful antibiotics discovered using AI

Machine learning spots molecules that work even against 'untreatable' strains of bacteria.

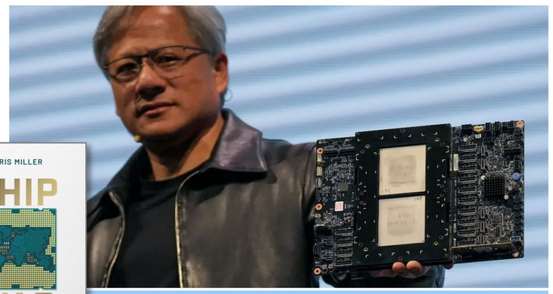
Article

Highly accurate protein structure prediction with AlphaFold

...and the engine powering AI is a GPU

If AI Is a Gold Rush, Nvidia Is Selling Shovels

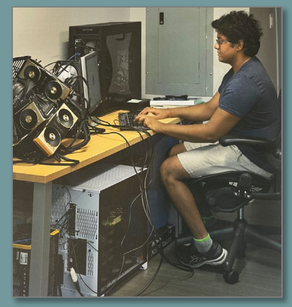
Without Nvidia's GPUs, there'd be no ChatGPT.



Input form researchers at the AI Pioneer Centre

"To execute our own research ideas, we are often forced to collaborate with foreign colleagues who have access to large GPU facilities"

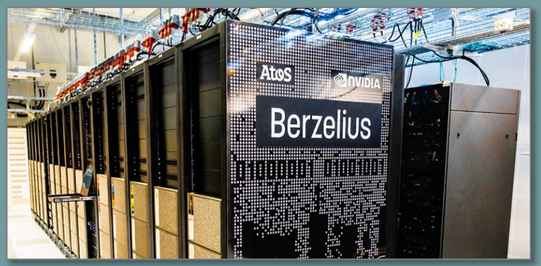
Strategy: "Go big" to accelerate DK!



Local approach

- Small, local, uncoordinated investments
- Impossible to do big AI projects in Denmark

VS



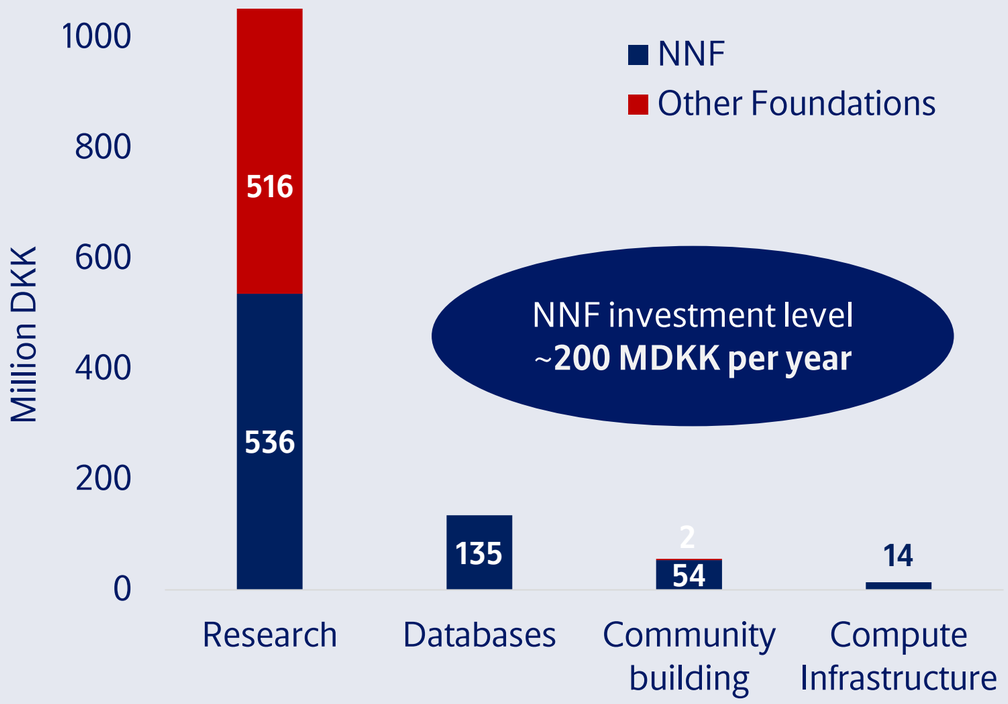
National Supercomputer

- Professional, sustainable, compute at scale

Motivation: why is NNF investing in an AI Supercomputer?

Danish investments in AI and data science

Graph: total sums granted 2020–2022



Wallenberg Foundation's investments in AI and data science

WASP – **Wallenberg AI, Autonomous Systems and Software Program**

- **Funding:** 4.9 billion SEK over 2015-2031 to research in AI, data science, computer science, math & algorithms
- **~200 MDKK per year invested in AI & digital research** (not infrastructure)

Data Driven Life Science Program

- **Funding:** 3.1 billion SEK from 2020-2032 to SciLife Lab and 10 universities
- **~180 MDKK per year invested in Data Science and AI for Life Science**

Berzelius – a national supercomputer for AI research

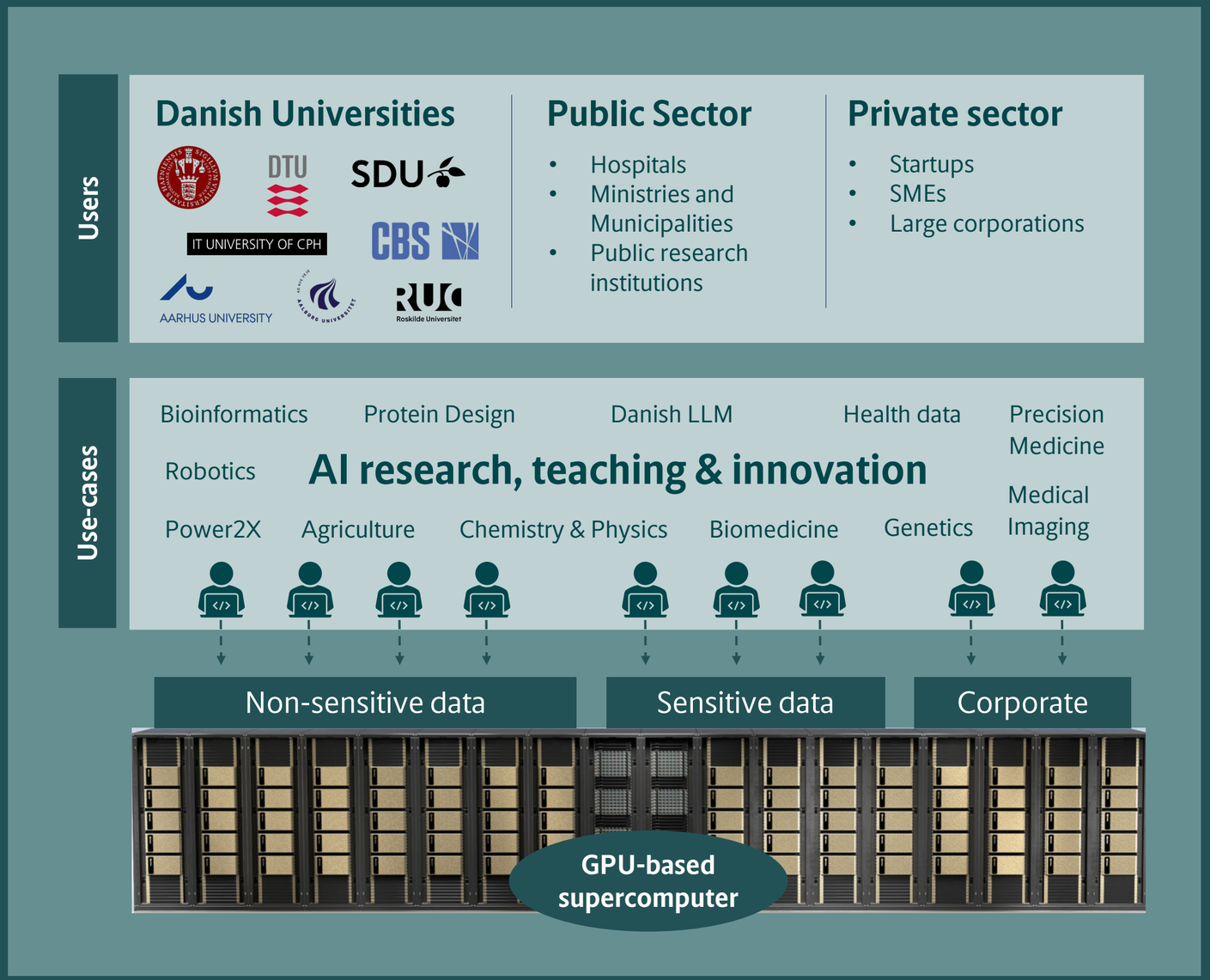
- **Funding:** 300 million SEK grant awarded in 2020
- **~250 MDKK total investment in GPUs**

Take-home message #1: Computing is not free

...in fact, it's an increasingly big expense in research and someone needs to pick up that bill

Remember to include computing expenses and data related expenses (storage, data management, etc.) in the budgets of each NNF funded project, both stand-alone and open competition

The operating model



Access depends on

- Danish affiliation
- Acceptance of charter (ethics, etc.)
- Access committee approval
- Payment

Payment models

- Full cost rate (e.g. for corporates)
- Reduced rate (e.g. for academia)
- Grant pools (e.g. NNF or other funders granting “free” compute)

For pilot stage

Pilot project criteria

- Mature projects ready to scale
- Team and data in place
- Diversity of use-cases and user types

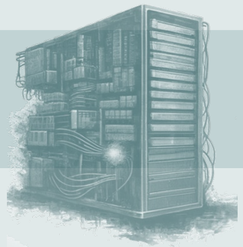
Got an idea for a pilot-phase job?
Contact me (mba@novo.dk)

Development of Gefion & DCAI

Mapping and scoping phase	Design and negotiation	Planning and implementation	Launch
Project core team NatTech Innovation	Project core team NatTech Innovation Portfolio Management Finance Legal	Project core team NNF Management Project Excellence External advisors (Tech/Science) Finance Legal NatTech Corporate affairs	Project core team NNF Management Project Excellence Finance Legal NatTech Corporate IT Corporate affairs Medical Science	**External advisors and contractors** • Tech: NVIDIA, EVIDEN, DK HPC experts, DeiC, PwC • Cybersecurity: PwC, NGC, SDU • Business models: IMPLEMENT • Legal: Bird&Bird, Gorissen • Recruitment: Korn-Ferry

Danish GPU compute Platform	Project Falcon	Project Gefion

Stages of Gefion



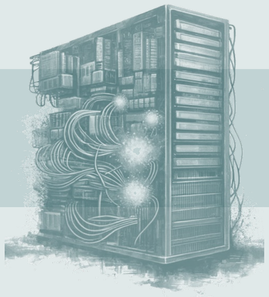
Gefion 2024

GEFION ON DANISH GROUND

State: HPC system as delivered by NVIDIA and EVIDEN. Focus is on building and testing the initial set of services.

Users: Developers, administrators and pilot users (friendly test-users)

Got an idea for a pilot-phase job?
Contact me (mba@novo.dk)



Gefion 2025

GEFION OPEN FOR USERS

State: Big shared GPU cluster fit for non-sensitive AI projects + a “private server” solution for major customers (e.g. corps).

Users: Academics working on non-sensitive data, start-ups, corporates, public sectors, etc.

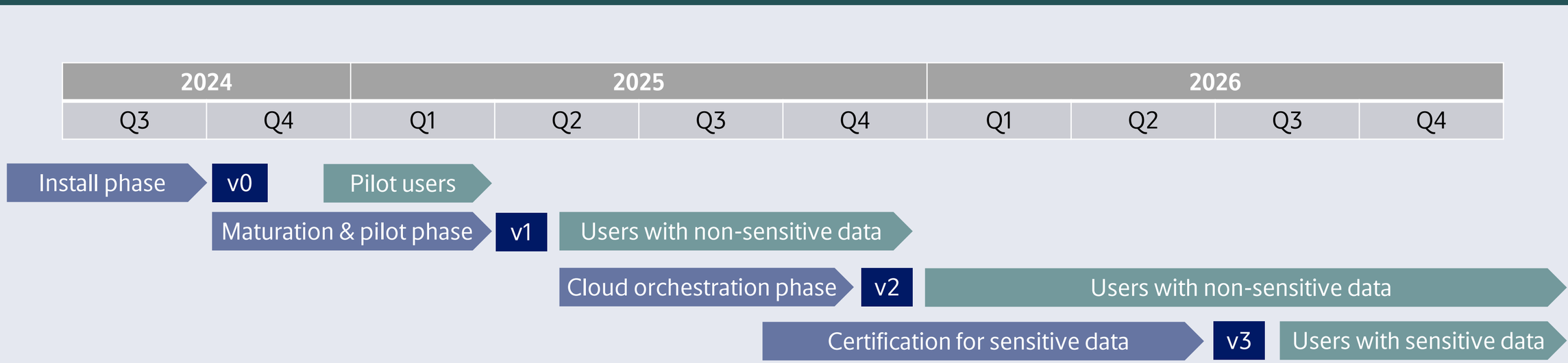
Gefion 2026+

CLOUD FOR SENSITIVE DATA

State: Secure, certified cloud-services that offer dynamic access to GPU computing at scale compliant for sensitive (health) data.

Users: all users, all use-cases

Stages of Gefion



Take-home message #2: Sensitive vs non-sensitive data

1. Biomedicine is a mix of non-sensitive and sensitive (think mouse/cell-line vs human). Gefion is relevant for these types of projects from day 1.

2. Clinical research is almost exclusively sensitive data. Gefion will become relevant for this around 2026.

3. Some types of data can be anonymized or otherwise be made less sensitive from a GDPR perspective (making Gefion relevant earlier for such projects).

Take-home message #3: It is much more than hardware

1. NVIDIA sees Gefion as an enabling tool for engaging with Danish researchers
2. NVIDIA will provide training, support, and code-optimisation
3. NVIDIA will engage with Danish researchers on solving the most pressing challenges in flagship projects
4. DCAI will be a central hub for accelerated GPU-based computing

NVIDIA and Danish collaboration flagships

Pharmaceuticals and biotechnology

- Drug development
- Protein design and protein LLMs

AI healthcare

- Human centric, democratic AI for healthcare
- LLM for clinical data

Acceleration of green transition

- Materials for energy technologies
- Data for digital twin of earth - remote sensing, earth data

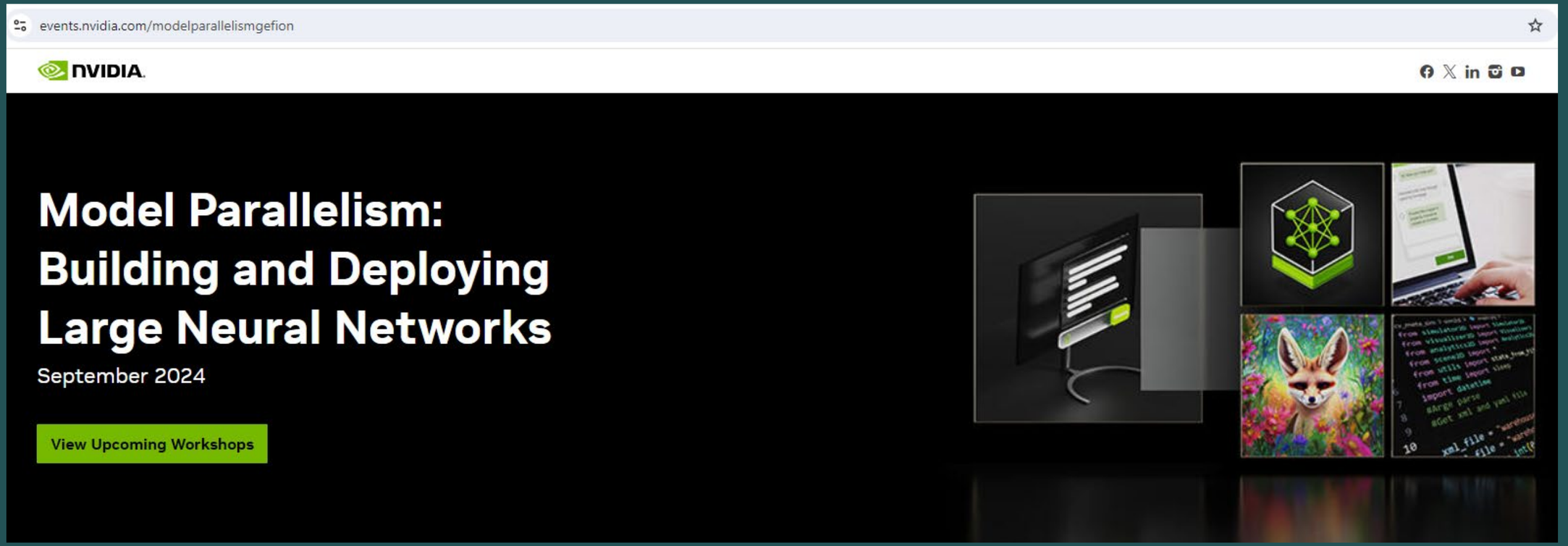
Fault tolerant quantum computing

- Simulation of quantum states
- Error correction schemes and integration of quantum computers with classical supercomputers

Fundamental AI research

- Explainable AI
- Removing biases in AI
- Causality and AI

Workshops and user training



CEO appointed last week!

Published

August 14th, 2024

The Novo Nordisk Foundation and the Export and Investment Fund of Denmark (EIFO) are pleased to announce the appointment of Nadia Carlsten as CEO of Danish Centre for AI Innovation A/S (DCAI), a company which will own and run Denmark's first AI supercomputer, Gefion.

Stay tuned...Gefion is on it's way!

Nadia

NNF contacts

Mette

Lasse

Morten

Ulrik

Follow the project @ https://novonordiskfonden.dk/en/projects/danish-centre-for-ai-innovation/