
Green Technology for sustainable AI

葉信和 / Hsin-Ho Yeh

Founder & CEO / Software Engineer @ 信誠金融科技

hsinho.yeh@footprint-ai.com



Download Slides

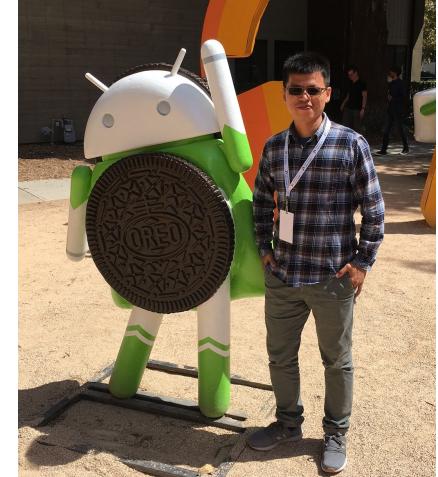
<https://bit.ly/47rBVs0>



[`https://github.com/FootprintAI/talks/tree/main/slides`](https://github.com/FootprintAI/talks/tree/main/slides)

About me

- 2020 - Present at 信誠金融科技
 - **Tintin**: a machine learning platform for everyone
 - <https://get-tintin.footprint-ai.com>
- 2016 - 2020 at IglooInsure (16M+ in series A+ 2020)
 - Provide digital insurance for e-economic world
 - Funded in KUL, Headquartered in Singapore
 - First employee/ Engineering Lead / Regional Head/ Chief Engineer
- 2013 - 2016 at Studio Engineering @ hTC
 - Principal Engineer on Cloud Infrastructure Team
- 2009 - 2012 at IIS @ Academia Sinica
 - Computer vision, pattern recognition, and data mining
- CS@CCU, CS@NCKU alumni





Intro & timeline

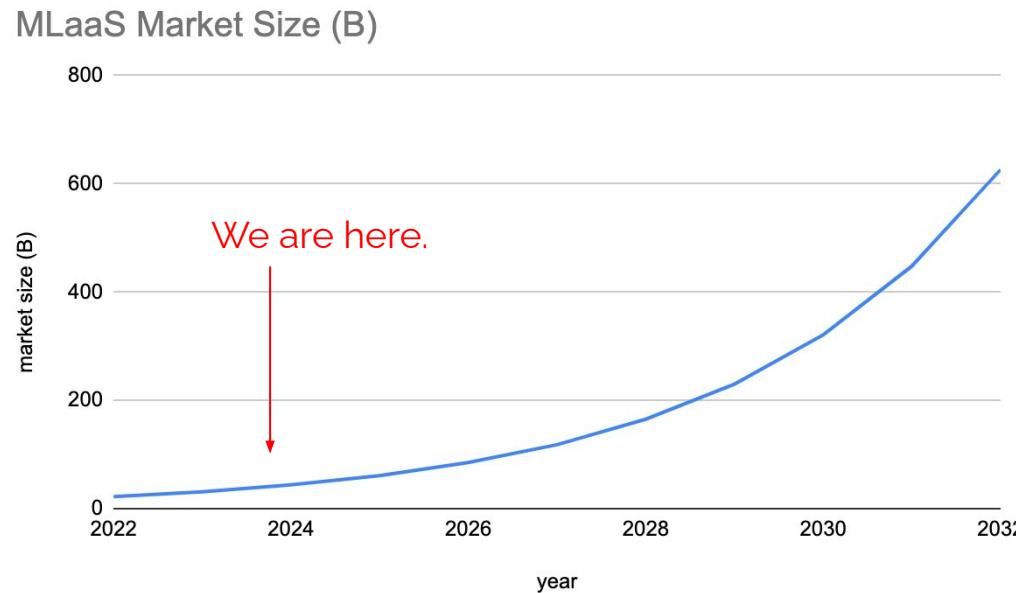
- footprint-ai.com (信誠金融科技) is committed to providing Software-as-a-Service for serving AI/ML applications, specializing in:
 - Cloud-native architecture, MLOps, Green Tech, Internet-scale Data Analysis
- Practical experience in taking an idea (zero) to many products (one)
 - Singapore-based insurtech startup iglooinsure (18M USD/2020 Series A, 46M USD/2023 Series B).
- Business partner with IBM Taiwan (IBM.tw)
- Joined Nvidia Inception program in 2021.
- Moved into FinTechSpaces in 2023.

Agenda

- Trends and observations.
- What is Sustainable AI?
- What is Green Technology / Green Software ?
- What is our effort on green economy?



MLaaS is the future

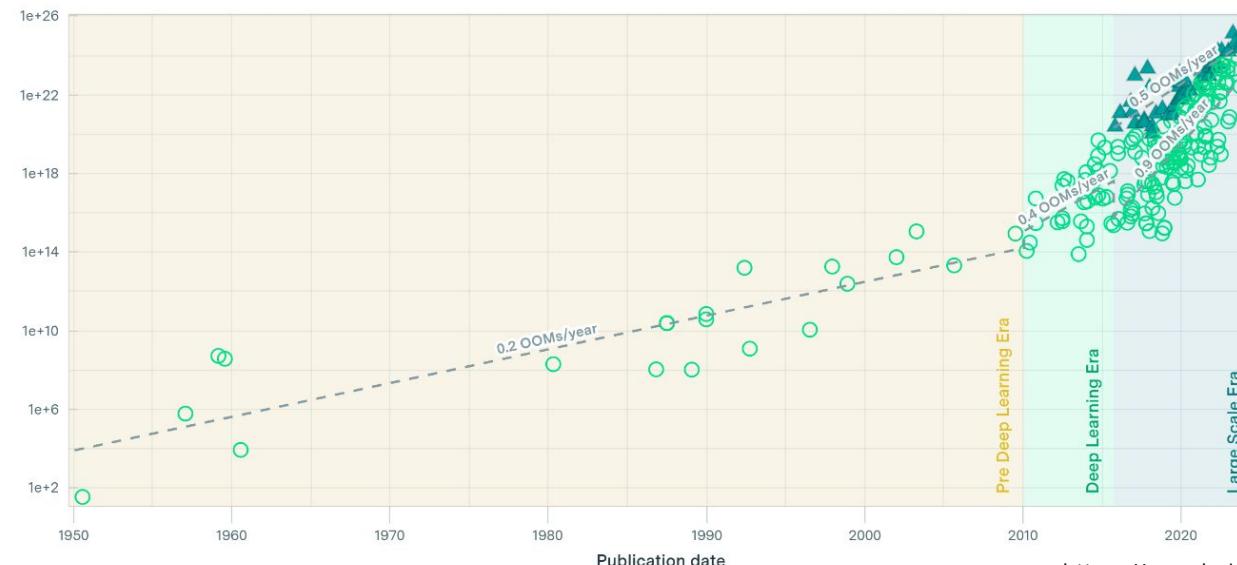


ML Model Growth history

Training Compute of Notable Machine Learning Systems Over Time

≡ EPOCH

Training compute (FLOP)



But the gas emission is also accelerating ...

Common carbon footprint benchmarks

in lbs of CO₂ equivalent

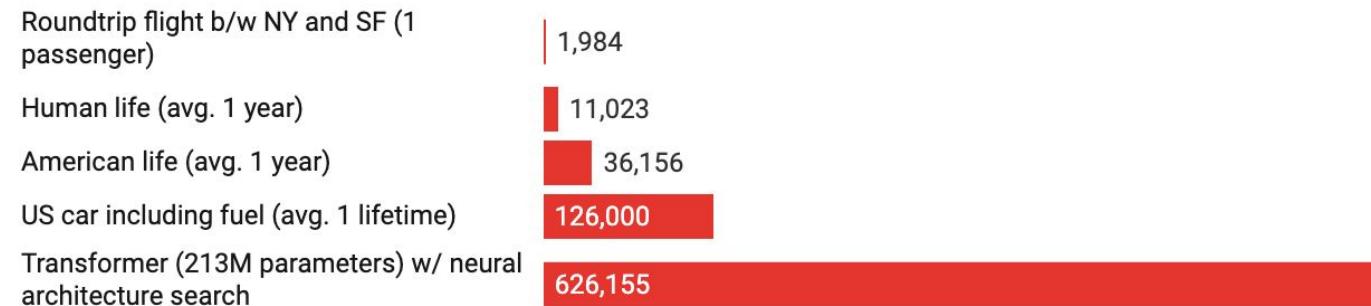
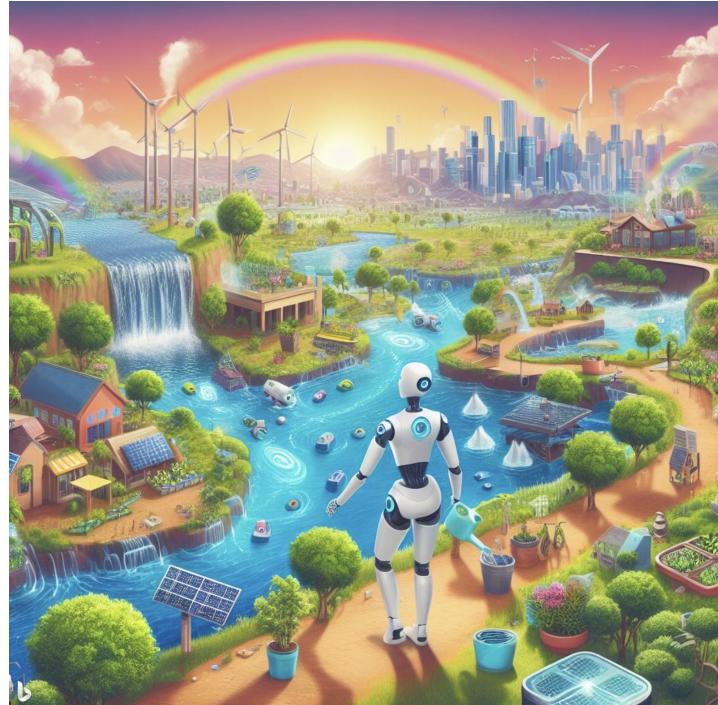


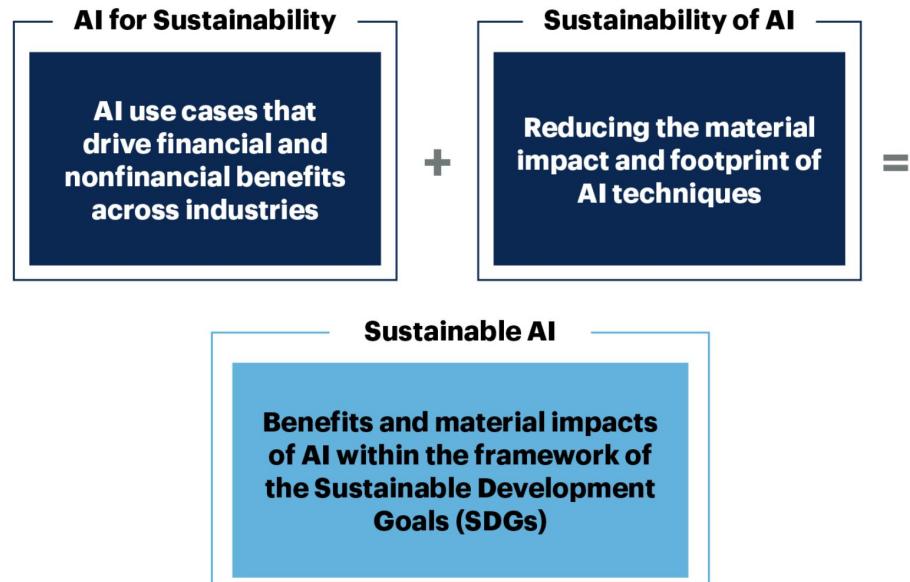
Chart: MIT Technology Review • Source: Strubell et al. • [Created with Datawrapper](#)



<https://www.bing.com/images/create/people-don27t-care-about-carbon-emission-and-water-/654fe02659aoa441e91d57dfd6b950bfa?id=JNCSWj7OssHkQcitT2C9yQ%3d%3d&view=detailv2&idpp=genimg&FORM=GCRIDP&mode=overlay>

<https://www.bing.com/images/create/create-a-photo-where-an-ai-is-embrace-sustainability/654fe07a3f8545e099187f40b34fded9?id=Y4ngLElWrhJD%2fMcqQJXcg%3d%3d&view=detailv2&idpp=genimg&FORM=GCRIDP&mode=overlay>

What is sustainable AI



gartner.com

Note: SDGs were designed by the United Nations General Assembly in 2015 as a call to action to pursue sustainable development.

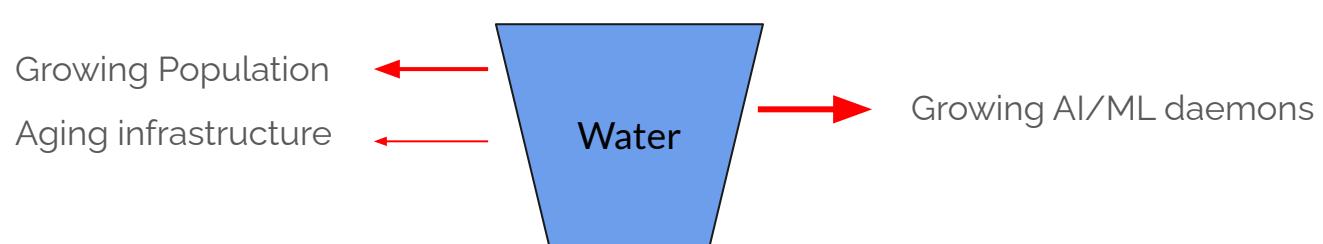
Source: Gartner
© 2023 Gartner, Inc. All rights reserved. CM_GTS_2477247

Gartner

What is sustainable AI

We need AI/ML to accelerate our daily tasks and automation, but ...

- Data Centers (DCs) in US are consumed **2%** electricity use.
- ChatGPT needs to “drink” a 500ml bottle of water for every simple 20-50 questions and answers (and GPT-4 is even thirstier) [1]



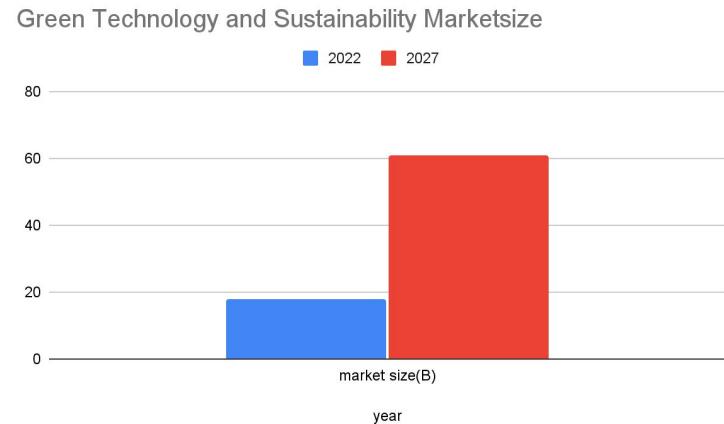
[1] <https://arxiv.org/pdf/2304.03271.pdf>

What is Green Technology

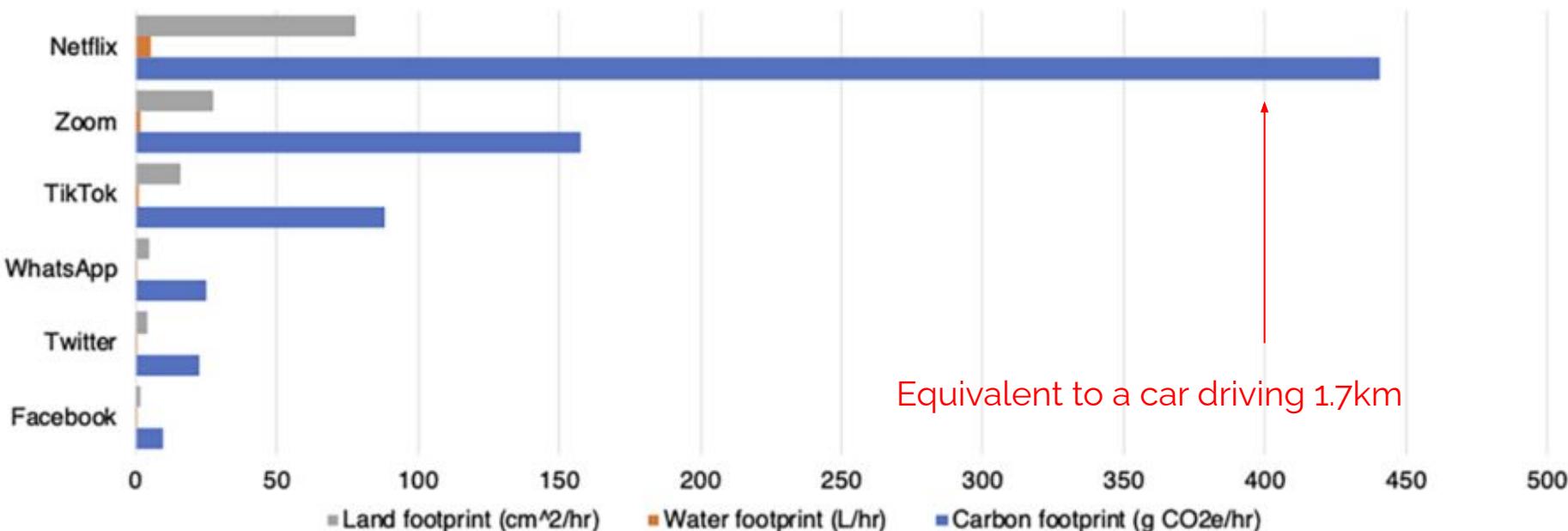
Green Technology = use of technology and science to reduce human impacts on the natural environment

Including:

- Alternative Energy
- Electric Vehicles
- Sustainable Agriculture
- Recycling
- Carbon Capture

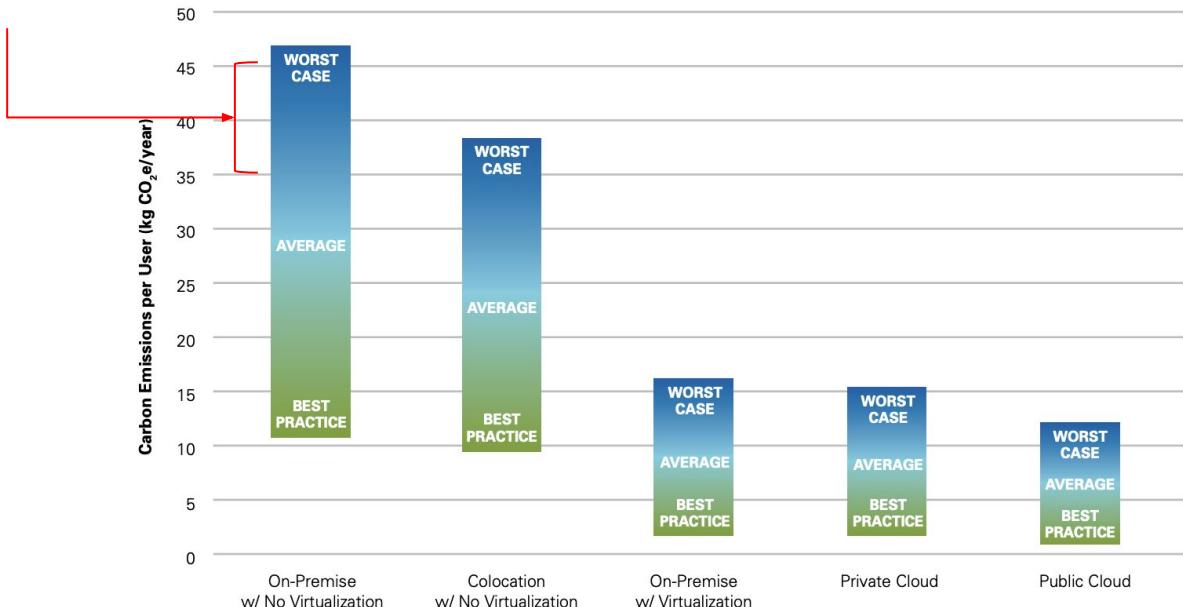


What is the environmental footprint of your app?

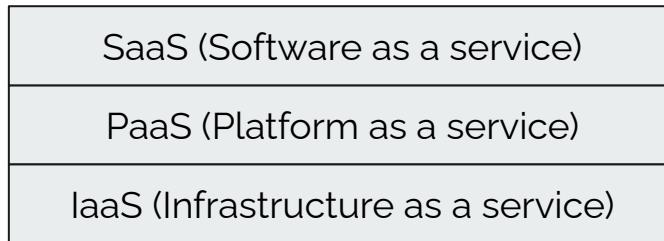


Carbon footprint: cloud vs on-premise

30% usage are on powering and cooling servers due to poor maintenance and practices.



Green Stack over the Cloud era



Azure Sustainability



Evolving beyond an operational approach

Sustainability is often viewed from an operational point of view, with the environmental mission treated as a cost center in the business, as a risk or compliance to be managed. But for Azure, datacenters and sustainable cloud infrastructure are more than just business—they provide an incredible opportunity to give back to the communities where we operate.

Azure has committed to focus on **four key areas of environmental impact to local communities—carbon, water, waste, and ecosystems.**

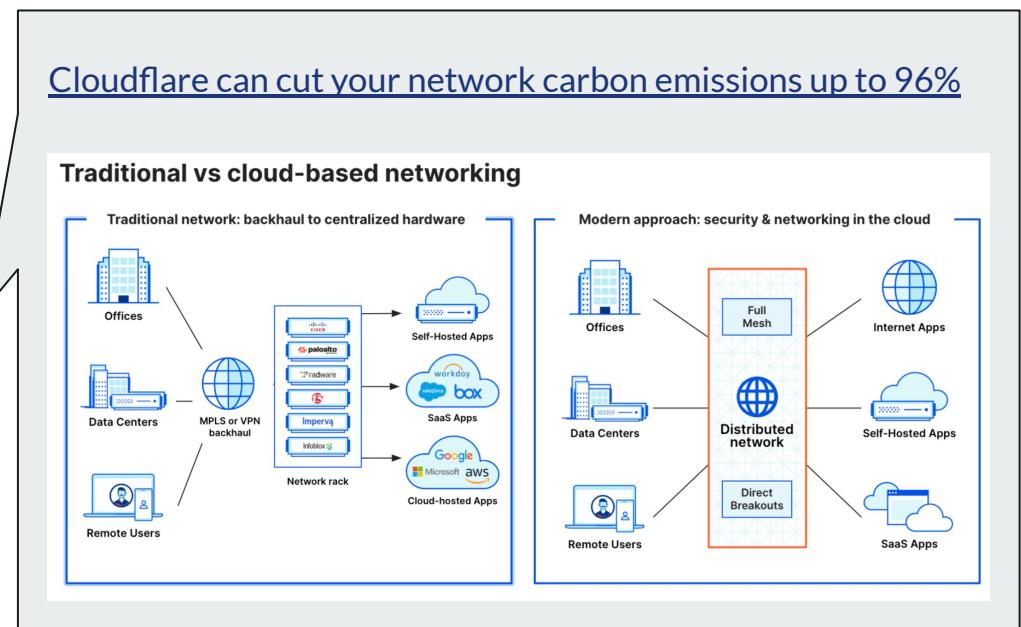
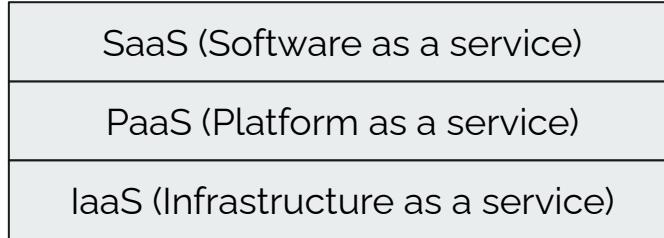
100% renewable energy by 2025.	Water positive by 2030 replenish more water than we consume by 2030.	Zero-waste certification by 2030.	Net-zero deforestation from new construction.
--	--	---	---

AWS Sustainability



5x AWS infrastructure is up to 5 times more energy efficient than typical European data centers	2.4 Billion Liters of water are returned to communities each year from replenishment projects completed or underway	90% In 2022, 90% of the electricity consumed by Amazon was attributable to renewable energy sources
---	---	---

Green Stack over the Cloud era



“The gains we get from hardware, in terms of output per megawatt, will be dwarfed by the gains we get from software”

**Paul Calleja,
director of research computing services
University of Cambridge**



<https://www.bing.com/images/create/the-gains-we-get-from-hardware2c-in-terms-of-output/65517760b1c24629aef5e094f85edb0?i=yO85O15ojMiiKQdM5WD62w%3d%3d&view=detailv2&idpp-genimg&FORM=GCRIDP&mode=overlay>

What is Green Software? (1/4)



Green software engineering minimizes environmental impact by integrating sustainable practices into software development, creating **responsible** systems that are **energy-efficient** and **resource-optimized**

<https://engineering.leanix.net/blog/sustainable-green-software-engineering/>

What is Green Software? (2/4)



- Architecture/design decision
 - **Static code analysis** for estimating carbon emission
 - Real time power consumption monitoring with **dynamic code analysis**.
 - **Serverless computing** shares infrastructure resources by executing functions only on demand.

What is Green Software? (3/4)

- Green Coding Practice
 - Choose statically-typed language over scripting language.

Time & Memory	Energy & Time	Energy & Memory	Energy & Time & Memory
C • Pascal • Go	C	C • Pascal	C • Pascal • Go
Rust • C++ • Fortran	Rust	Rust • C++ • Fortran • Go	Rust • C++ • Fortran
Ada	C++	Ada	Ada
Java • Chapel • Lisp • Ocaml	Ada	Java • Chapel • Lisp	Java • Chapel • Lisp • Ocaml
Haskell • C#	Java	OCaml • Swift • Haskell	Swift • Haskell • C#
Swift • PHP	Pascal • Chapel	C# • PHP	Dart • F# • Racket • Hack • PHP
F# • Racket • Hack • Python	Lisp • Ocaml • Go	Dart • F# • Racket • Hack • Python	JavaScript • Ruby • Python
JavaScript • Ruby	Fortran • Haskell • C#	JavaScript • Ruby	TypeScript • Erlang
Dart • TypeScript • Erlang	Swift	TypeScript	Lua • JRuby • Perl
JRuby • Perl	Dart • F#	Erlang • Lua • Perl	
Lua	JavaScript	JRuby	

What is Green Software? (4/4)



- Moving to **Cloud** or **Edge** Computing.
 - Approaches like virtualization, energy-efficient data centres, multi-tenancy, etc., enable cloud computing to reduce energy usage and carbon emissions.
 - Edge computing allows you to redistribute computation near the users to lower costs
- Building the mindset
 - Trade-off between business and environmental goals.

Kafeido: machine learning platform for green economy

Machine Learning Platform For Green Economy

Kafeido : Machine Learning Platform For Green Economy

Our one-step platform equips both the serverless and the exclusive micro-model deployment architecture to achieve real-time machine learning model deployment for the green environment. Kafeido can not only save your personnel costs on model operations but also avoid the excessive expenses on hardware and electricity.

Features highlighted

- POINT 01** Kafeido is Your Best Choice
- POINT 02** Developing with heterogeneous frameworks
- POINT 03** Serverless Architecture and Horizontal Expansion Advantages
- POINT 04** Micro-model Green Deployment Architecture

How does Kafeido work?

Applicable places: community, school, hospital, shopping mall, factory, corporate
Deployment plan: On-prem, SaaS

- Step 01** Select a model from our model zoo or upload your own model
- Step 02** Generate the model deployment
- Step 03** Select your data source
- Step 04** Automatic management model inference

Customer success story : Sustainable Smart City Monitoring Center

- Existing Challenges**
The traditional command and control center (or called ICCO) simply collects all information from each camera to one control panel. The security personnel can monitor the entire city conveniently and provide adequate assistance as needed. However, when the monitoring scope increases (e.g. from CCT to drone shooting, from single point to field monitoring...etc.), long-term dependence on the security personnel not only increases personnel costs dramatically but also fails to operate in high quality.
- ESG Awareness**
Environmental monitoring (e.g. air pollution, factory waste gas emission...etc.) is another challenge—how to import the green machine learning platform to keep a city green while minimizing carbon emissions also becomes a hot topic.
- Kafeido Accomplished**
Combining multiple data sources from drones and CCTV cameras and various machine learning models, Kafeido triggers warning events based on the model inference results to notify related personnel via SMS and/or email and achieves a 24/7 decentralized monitoring process. With Kafeido, you can easily scale out your data sources for monitoring and state-of-the-art machine learning models to guard valuable properties for you and for our next generation.

Our professional software architecture and green-oriented solution help your business apply much easier and more affordable AI technologies!

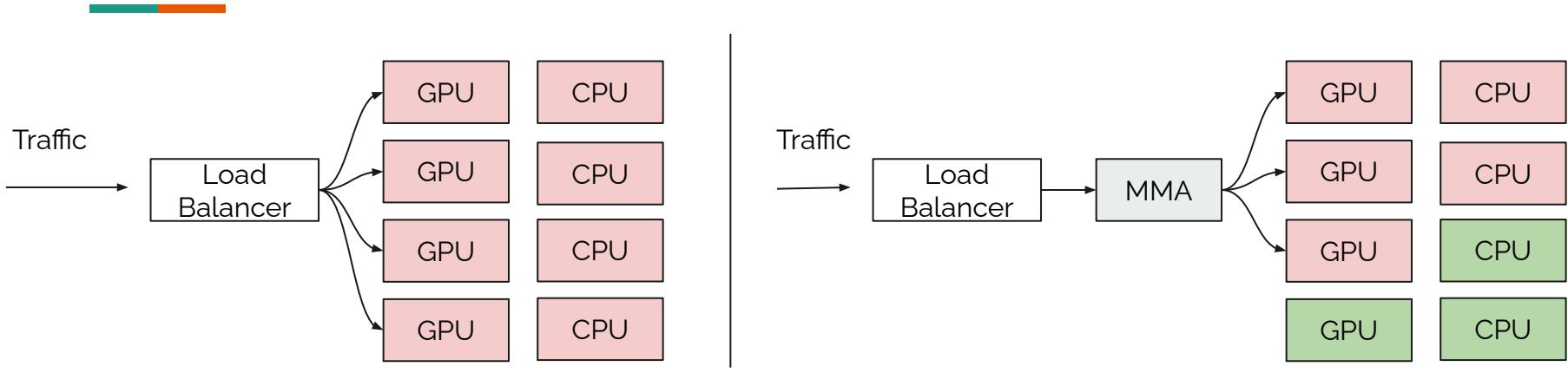
Contact Us
信誠金融科技股份有限公司 (footprint-ai.com) dedicates in developing machine learning platform and providing AI-oriented software services. We are expert in machine learning platform, data middle platform, and their customizations.
Address : No. 287-2, Sec. 3, Chengde Rd, Taipei City 103, Taiwan.
Email : kafeido@footprint-ai.com

Tintin
Machine Learning Platform For Everyone

NVIDIA.
INCEPTION PROGRAM

XINCHENG FINTECH CO., LTD.

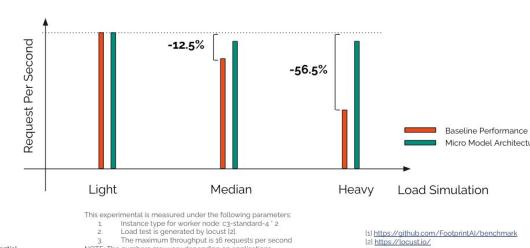
Increasing utilization with micro model architecture



Energy Cost Efficiency - GPU Inference Case Study



Energy Cost Efficiency - CPU Inference Case Study



<https://kafido.com/footprintai/benchmark>

<https://kafido.ai>

<https://kafido.com>

© 2023 Xinchen fintech Co. LTD Confidential.

Product Comparison

Product	kafeido	https://granulate.io/
via	Horizontal pod autoscaler	Vertical pod autoscaler (continue profiling and optimization on resource usage)
Installation	pre-compile	profiler
Optimization	Runtime	Runtime (dynamic adjustment resource request/limits)
Tech stack	kubernetes	kubernetes
Valuation	?	650M (acquired by Intel) 2020/12M/A, 2021/30M/B
用吃的來比喻	美食街(多店家共享餐桌)	夜市攤販(生意多圍的地就廣)

Application I - Smart Surveillance Solution for Green City [2022Q4, launched]



Parametric Insurance

For protect uncertainty when we move to sustainable solution.



<https://www.bing.com/images/create/parametric-insurance-for-protect-uncertainty-when-/654ff2b8b6f148aaaae103722a52b9b7?id=MwAEAGt%2bwrmf4Kn5vgEpaPg%3d%3d&view=detailv2&idpp=genimg&FORM=GCRIDP&mode=overlay>

Parametric Insurance

- One variable insurance
 - Agriculture insurance where the farmer wants to protect his/her yields from rainfall over a time period
- Payout is implemented when a certain event happened in a certain time period
- I.e. $f(\text{time period}) > \text{threshold}$, where f could be
 - Weather condition
 - Bitcoin pricing
 - Cloud service status



Existing Parametric Insurance

- Weather-based index insurance
 - For protecting farmers loss due to extreme climate change.
- Renewable Energy Insurance
 - For mitigating construction and operation risk, such as Offshore wind power and photovoltaic.
- Energy Efficiency Insurance
 - For easing risks from business interruption, material damage, and so on.

Parametric Insurance for NEXT

- We are designing a new parametric insurance for guarding our virtual properties.
-
- Feel free to reach out if you have idea and **We are hiring.**



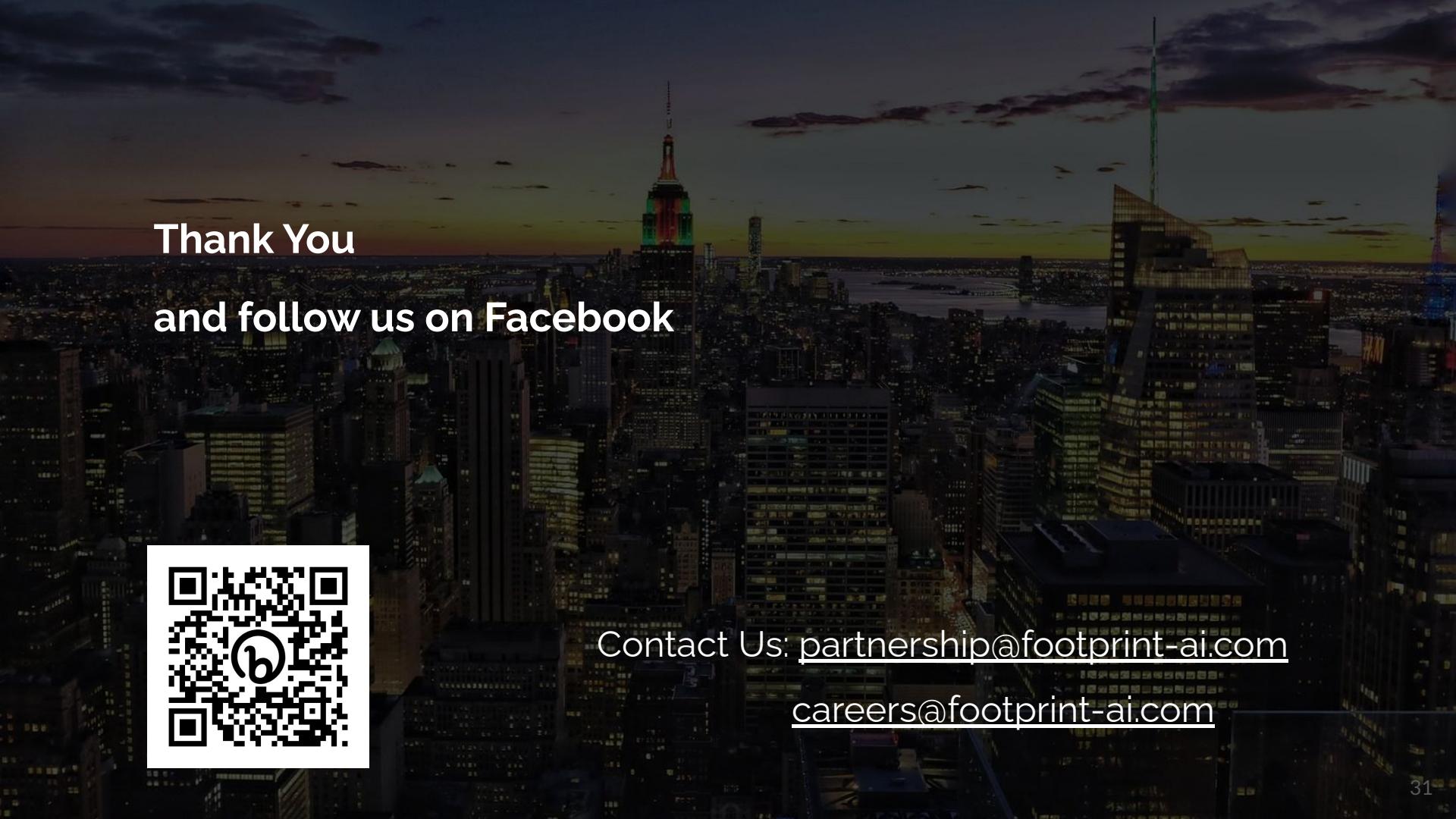
<https://www.bing.com/images/create/c...?id=vYRditgOsNkH6GtTwudwFg%3d%3d&view=detailv2&idpp=genimg&FORM=GCRIDP&mode=overlay>



Key take a way

- Software, not hardware, plays a key role the way of combating and preserving environment.
- Choose cloud over on-premise solution and adopt newer technologies for better conserving environment.

And this LONG journey has only just begun...



Thank You

and follow us on Facebook



Contact Us: partnership@footprint-ai.com

careers@footprint-ai.com