Predicting Sustainable High Performance Computing

Dejan Milojicic, HPE Fellow and VP, Hewlett Packard Labs

ModSim 2024, Seattle, WA

IEEE
FUTURE DIRECTIONS

Technology Megatrends

Industry Advisory Board (IAB) of IEEE Future Directions Committee (FDC)

Metin Akay, Saba Al-Rubaye, Priscilla Amalraj,
Ravikiran Annaswamy, Jyotika Athavale, Klaus BEETZ,
Nuno Borges Carvalho, Kirk Bresniker, Valerie Browning,
Hong Chen, Tom Coughlin, Celia Desmond, Stephen Dukes,
Izzat El Hajj, Eitan Frachtenberg, Jean-Luc Gaudiot,
Shashank Gaur, Gustavo Giannattasio, Chris Gorog, Eric Grigorian,
Kathy Grise, Michael Gschwind, Mazdak Hashemi, Mike Ignatowski,
Charlie Jackson, Lizy John, Mrinal Karvir, Steve Keckler,
Witold Kinsner, Bruce Kraemer, Rakesh Kumar, Luis Kun,
Phil Laplante, Tim Lee, Maike Luiken, Deepak Mathur,
Dejan Milojicic (chair), Chris Miyachi, Paul Nikolich, Damir Novosel,
Sudeep Pasricha, Nita Patel, Liliane Peters, Sohaib Qamar Sheikh,
Jeewika Ranaweera, Roberto Saracco, Vesna Sossi,
George K. Thiruvathukal, William Tonti, John Verboncoeur,
May Wang, Rod Waterhouse, Stefano Zanero, and George Zissis.

31st March 2024

Introduction

- Megatrends influence humanity in many ways
- Technology megatrends are intertwined with economic, ecological & social megatrends
- The IEEE FDC IAB members determined the following three technology megatrends
  - Digital Transformation; Sustainability; and Artificial General Intelligence (AGI)
- Because megatrends may evolve over a 20 year or longer timeframe, this report describes an ensemble of technologies within these three megatrends
- We provide insights about technologies and megatrends and their impact on humanity
- We compare our insights with those of the IEEE Computer Society and position our predictions with those of Google Trends, IEEE Xplore and US Patents intellectual property

What Constitutes a Megatrend?

- A megatrend has an impact on the evolution of multiple trends, hence the importance to understand Megatrends
  - it is both the sum of individual trends and a guiding force since usually it leads to a perception that influences its components

- A megatrend impacts multiple factors, substantially
  - technological
  - economical
  - social
  - ecological

- Megatrend **is not**
  - temporary fashionable technology
  - coming from a single technical focus
  - of interest to a limited region or a group

- A megatrend **is**
  - of global, world-wide importance → Political
  - critical enough that will require regulation
  - encompassing multiple technologies
  - evolving over a few years if not decades

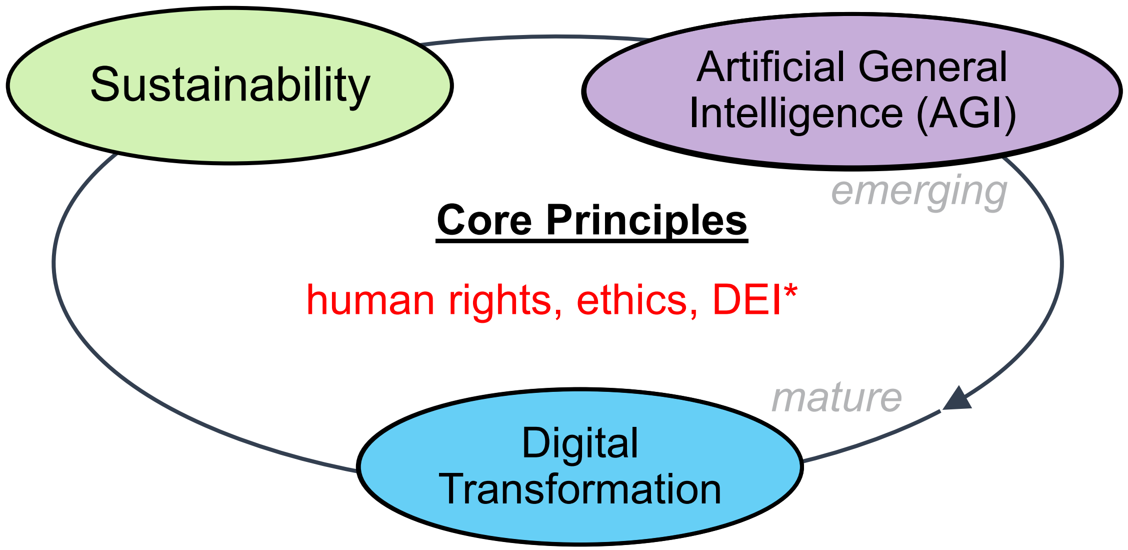
Portfolio of Predictions

- IEEE Future Directions Megatrends (THIS REPORT)
- Annual IEEE Computer Society Tech Predictions (Jan) and Scorecard (Dec), taking place for 15 years, since 2010
- Five special issues of IEEE Computer (2024, 2023, 2022, 2021, 2019)
- IEEE Computer "Predictions" Column (…. Jan'23, Apr'23, Jul'23, Oct'23, Jan'24, Apr'24, Jul'27)
- IEEE SCVS Industry Spotlights (Megatrends, AI, Sustainability, Digital Twins), co-sponsored by FDC, IEEE CS, IEC
- Special Features
  - IEEE SSE, "The Art of Prediction"
  - IEEE Design and Test, "Ethics in Sustainability"
  - IT Professional "What Gets You Hired Now Will Not Get You Hired Then"
- Many webinars, podcasts, keynotes, invited talks, panels, etc.
  - E.g. SXSW panel: "AI: Prosperity or Doom for Human Workforce?"
- Course "High Performance Computing: Use of AI and Emerging Technologies in Science"
- Decadal reports: Computer Society Report 2022 (issued in 2015); Future of Workforce (issued in 2023)

Process

- Team
  - We formed the team of approximately fifty people who meet throughout the year
  - Diversity
    - GEOGRAPHICAL: We have incorporated perspectives from the Middle East, Australia, Asia, Europe, and Latin America to US representation
    - GENDER: We have sixteen women out of fifty-four team members
    - TECHNICAL FIELD OF INTEREST: We have members from across 47 IEEE technical fields of interest
- The process and criteria are similar to IEEE CS Technology Predictions process
  - Selection of megatrends and associated technologies
    - During the inaugural year of 2023, we identified 3 megatrends: digital transformation, sustainability, and artificial general intelligence
    - For each megatrend, the team proposed approximately twenty technologies per megatrend
    - This was followed by down-selection to six technologies per megatrend, having each member at the time vote
  - Criteria and grading scale used by the team members for predictions
    - (A-F) for: Predicted Technology Success in 2023; (Potential for) Impact to Humanity; Predicted Maturity in 2023; Predicted Market Adoption in 2023
    - (1 year, 3y, 5y, 10y, 15y) Horizon view to Commercial Adoption
  - Outcome of the process
    - Impact to humanity as a function of technology advancement, qualified by maturity, market adoption and time-to-adoption
    - We calculate and report our confidence levels as the standard deviation in voting, and bias as a correlation between individual grades
  - Qualifying outcomes
    - We conclude with our insights derived from opportunities

Technology Megatrends



**DEI: Diversity, Equity and Inclusion*

Technology- vs General-Megatrends

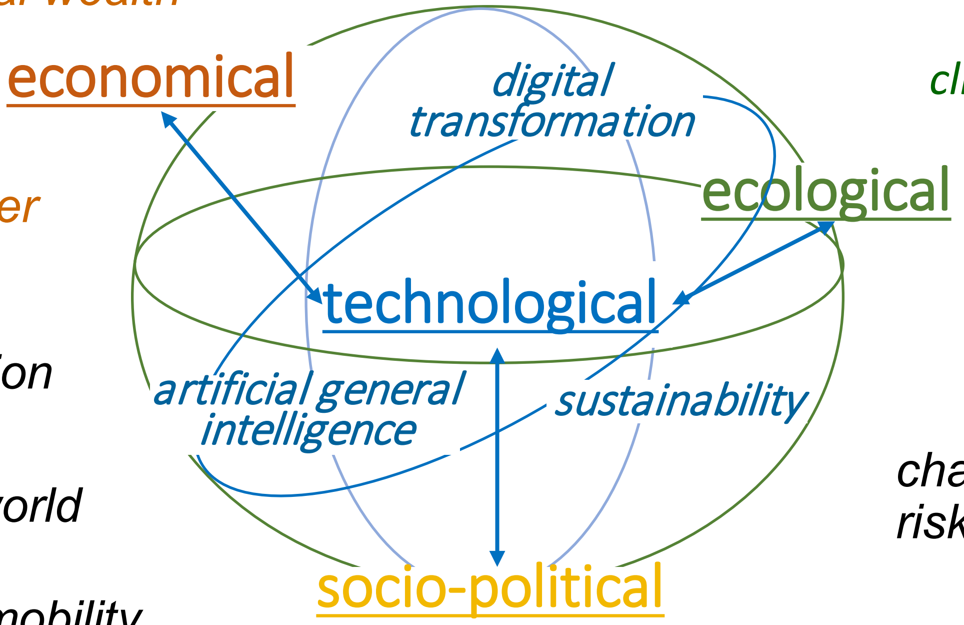
emerging markets

cost of climate change

increasing environmental pollution

emerging global wealth

climate change and resource scarcity



economical

digital transformation

ecological

technological

artificial general intelligence

sustainability

socio-political

shifting economic power

biotech/synthetic biology

globalization

changing disease burdens,
risk of pandemics

increasingly multipolar world

mobility

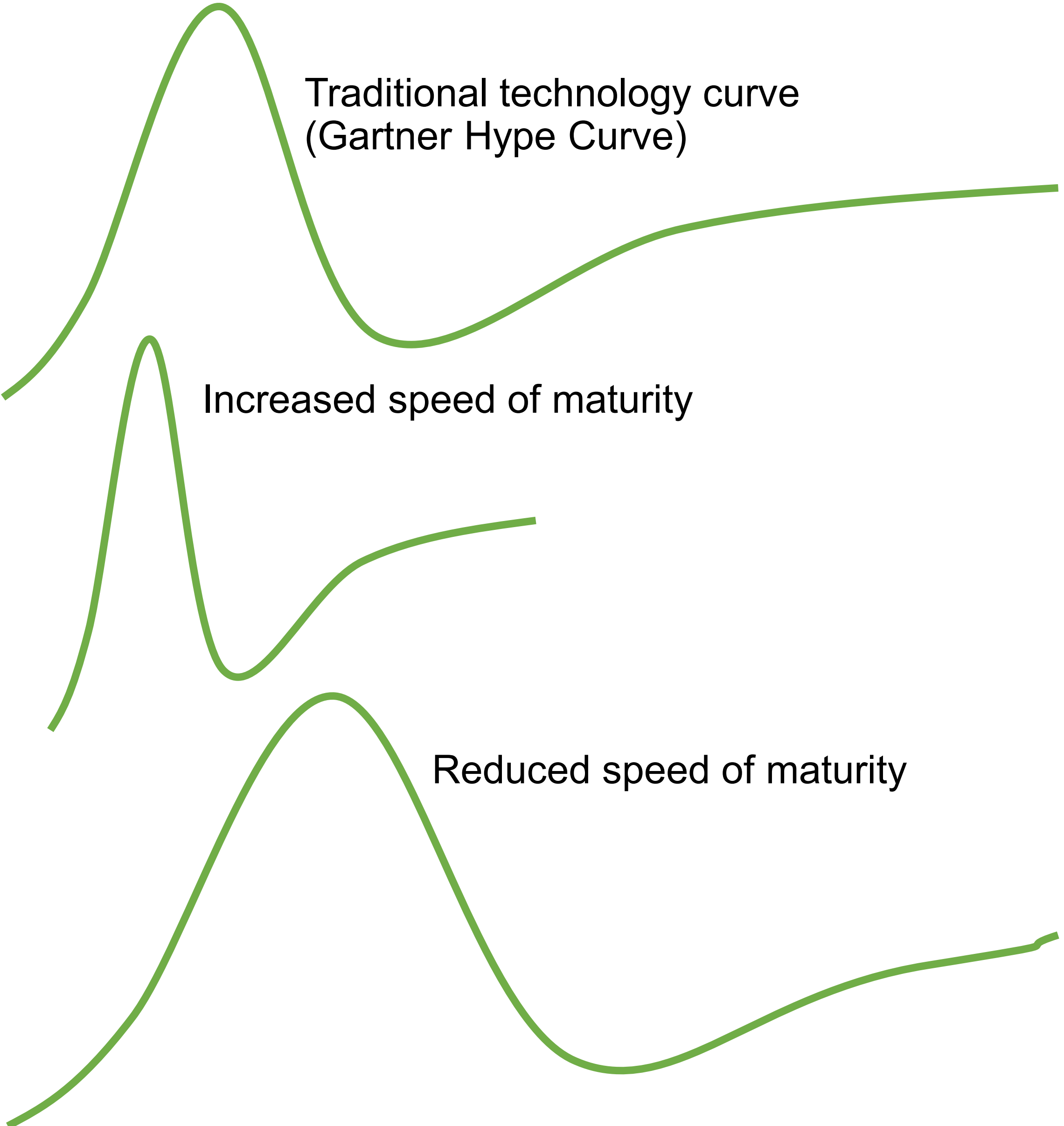
rapid urbanization

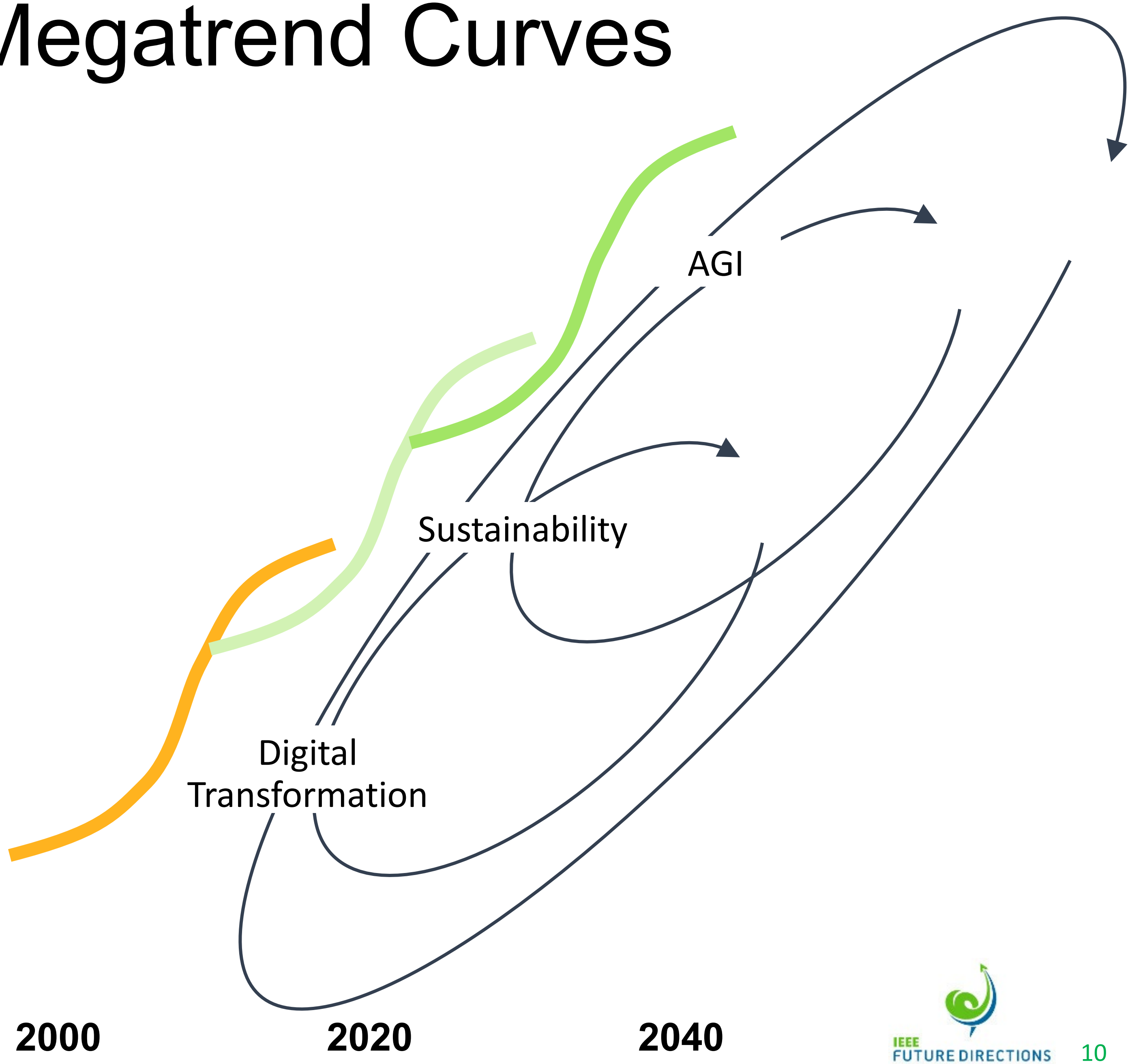
demographics and social change

aging population health

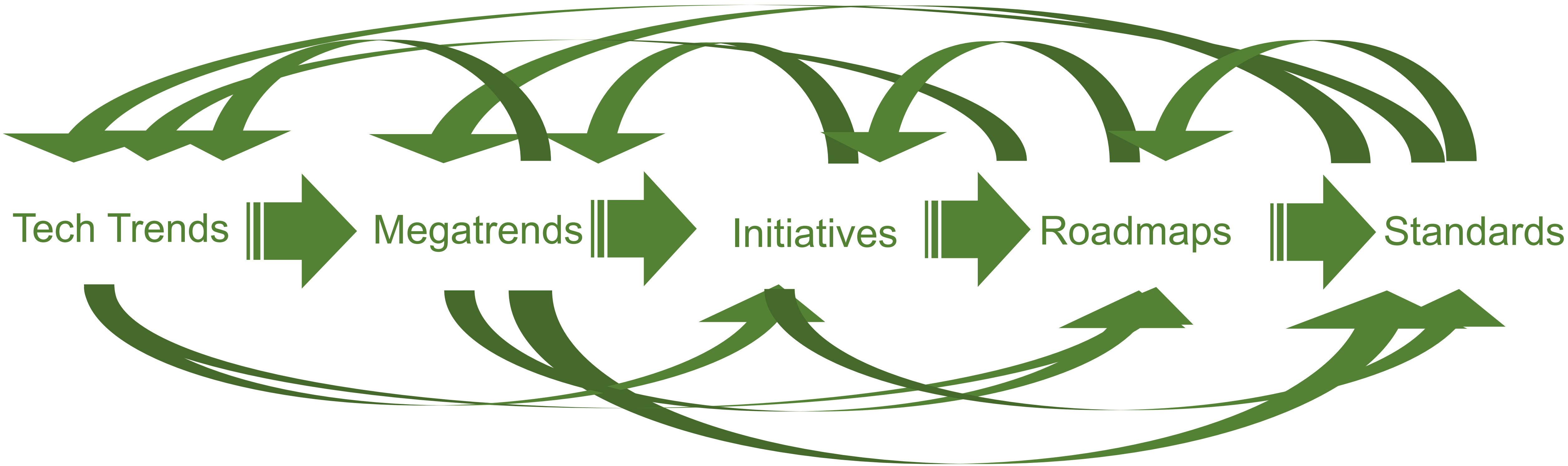
This Figure was originally published in C. Bash, K. Bresniker, P. Faraboschi, T. Jarnigan, D. Milojicic and P. Wood, "Ethics in Sustainability," in IEEE Design & Test, vol. 41, no. 1, pp. 25-32, Feb. 2024

Technology Trends vs Megatrend Curves

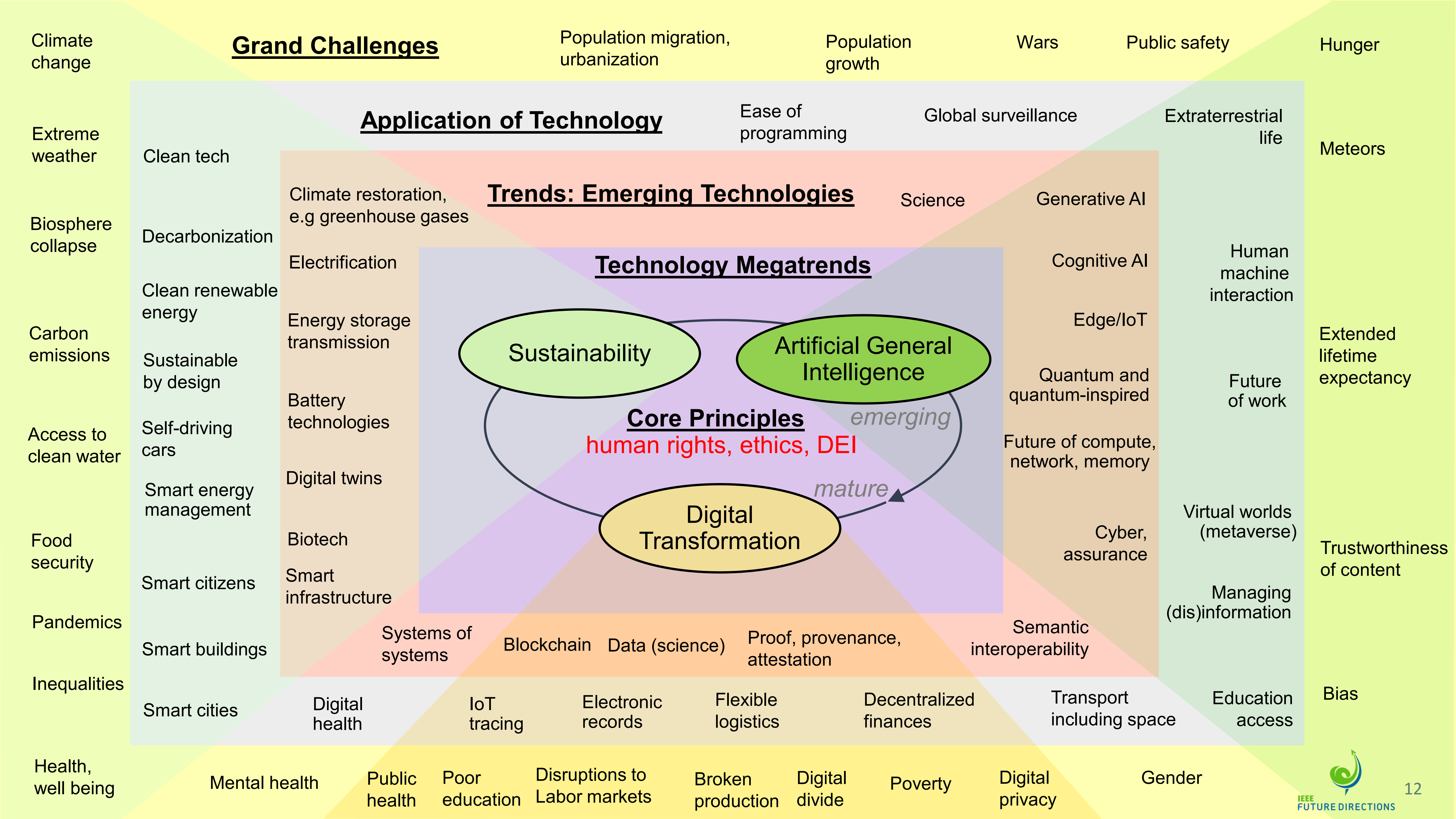

Traditional technology curve
(Gartner Hype Curve)
Increased speed of maturity
Reduced speed of maturity



Trends in the Broader IEEE Context



- Technology trends collectively result in observations about megatrends
- Megatrends help formulate and inform important IEEE Future Directions Initiatives
- Some successful IEEE Future Direction Initiatives result in IEEE Roadmaps
- Some trends, megatrends, initiatives, and roadmaps lead to industry standards



Grand Challenges

Climate change

Population migration, urbanization

Population growth

Wars

Public safety

Hunger

Extreme weather

Meteors

Biosphere collapse

Carbon emissions

Extended lifetime expectancy

Access to clean water

Food security

Trustworthiness of content

Pandemics

Inequalities

Bias

Health, well being

Mental health

Public health

Poor education

Disruptions to Labor markets

Broken production

Digital divide

Poverty

Digital privacy

Gender

Application of Technology

Ease of programming

Global surveillance

Extraterrestrial life

Clean tech

Decarbonization

Clean renewable energy

Sustainable by design

Self-driving cars

Smart energy management

Smart citizens

Smart buildings

Smart cities

Human machine interaction

Future of work

Virtual worlds (metaverse)

Managing (dis)information

Digital health

IoT tracing

Electronic records

Flexible logistics

Decentralized finances

Transport including space

Education access

Trends: Emerging Technologies

Climate restoration, e.g greenhouse gases

Science

Generative AI

Electrification

Energy storage transmission

Battery technologies

Digital twins

Biotech

Smart infrastructure

Cognitive AI

Edge/IoT

Quantum and quantum-inspired

Future of compute, network, memory

Cyber, assurance

Systems of systems

Blockchain

Data (science)

Proof, provenance, attestation

Semantic interoperability

Technology Megatrends

Sustainability

Artificial General Intelligence

emerging

Core Principles

human rights, ethics, DEI

mature

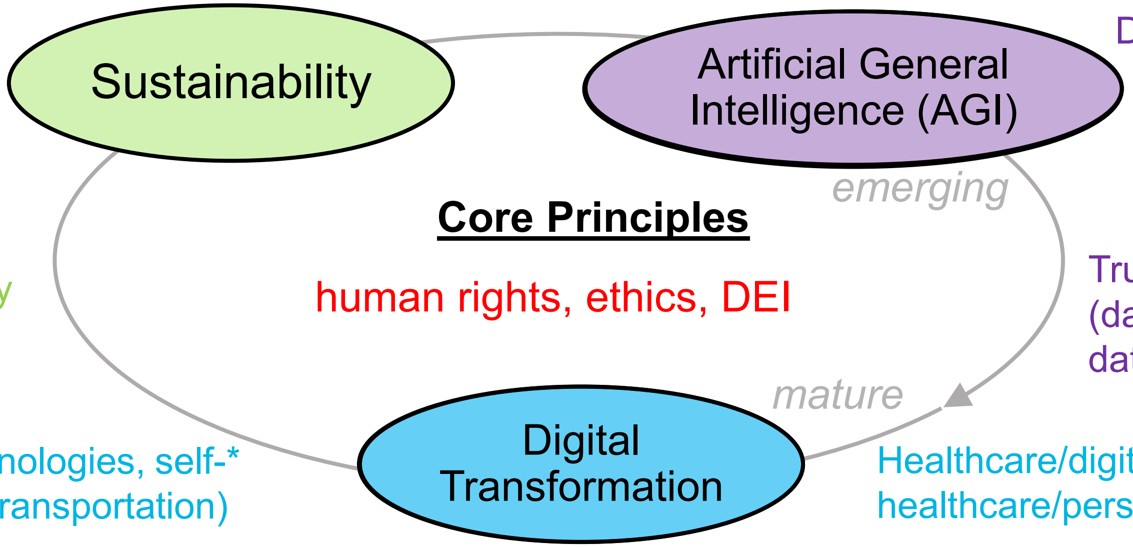
Digital Transformation

Relationship Between Megatrends

		How megatrend benefits		
		Digital Transformation	**Sustainability**	**AGI**
How megatrend contributes	**Digital Transformation**		• More control points • Clear separation and models • Opportunity to automate	• Broader set of applications • Edge-to-Cloud integration • Increases confidence
	Sustainability	• More incentives to transform • Reduced energy cost of transformation		• More powerful AGI • Broader adoption • Stretching limits
	AGI	• More effective transform • New ways of transform	• Innovating efficiency improvements • Improved anomaly detection	

This Table was modified from the table that originally appeared in P. Faraboschi, E. Frachtenberg, P. Laplante, D. Milojicic and R. Saracco, "Artificial General Intelligence: Humanity's Downturn or Unlimited Prosperity," in Computer, vol. 56, no. 10, pp. 93-101, Oct. 2023,

Megatrends Technologies



Sustainability

Artificial General Intelligence (AGI)

Digital Transformation

emerging

mature

Core Principles

human rights, ethics, DEI

Electrification

Sustainable IT and
IT for sustainability

Sustainable manufacturing,
recycling/Re-use

Energy storage

Decarbonization

Clean renewable energy

Generative AI, ChatGPT and other
application domains of AGI

Autonomous robots

Wearable and implantable technologies,
and computer brain I/F

Digital twins

Space technologies

Trusted computing
(data protection/privacy/cybersecurity/
data verification, lineage, etc.)

Healthcare/digital health/remote
healthcare/personalized medicine

Genomics/gene repair/DNA
medicine/molecular medicine

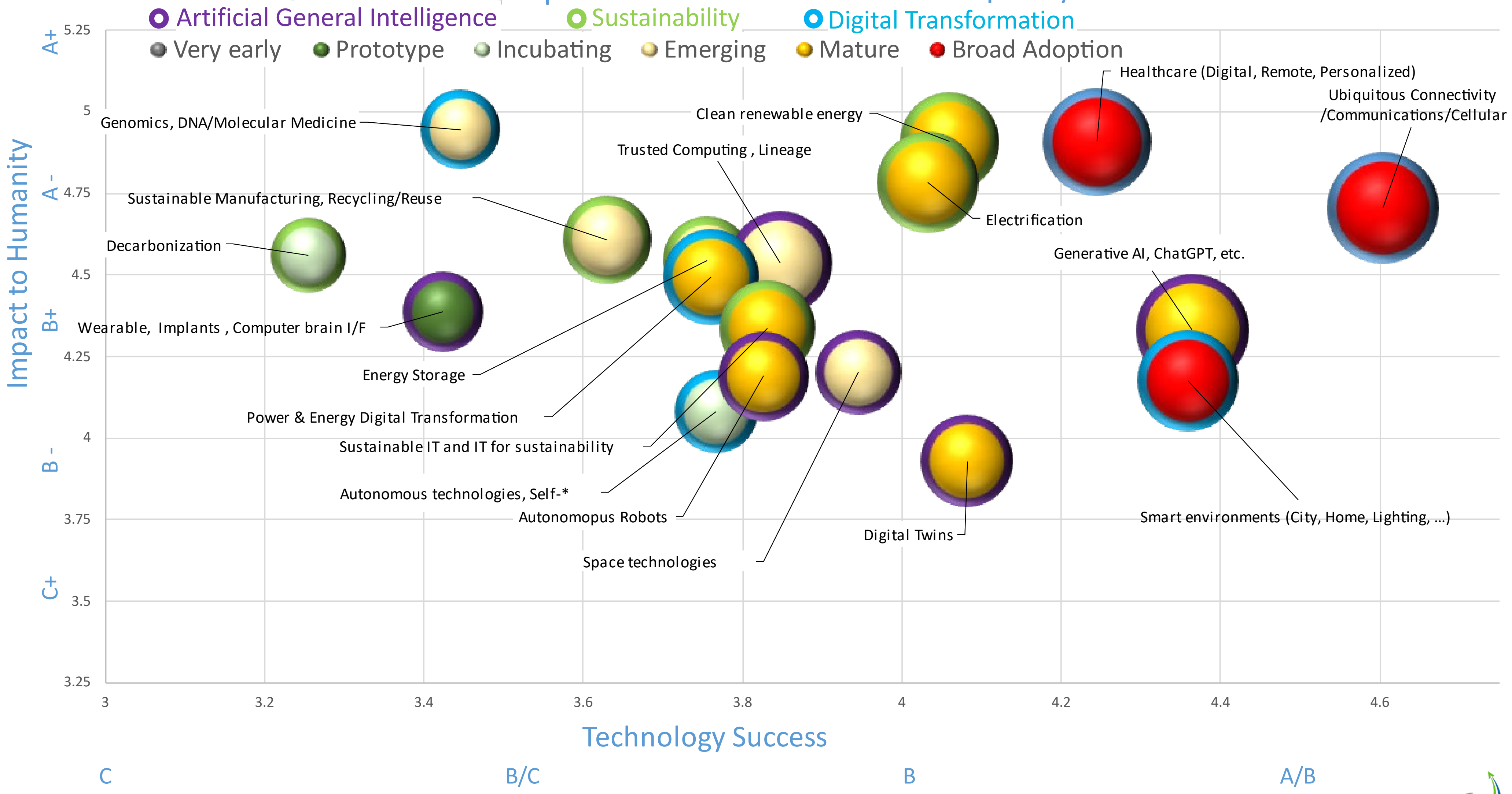
Power & energy
digital transformation

Ubiquitous connectivity/
communications/cellular

Smart environments
(city, home, lighting, …)

Autonomous technologies, self-*
(e.g. Self-driving transportation)

Megatrends to Technologies Mapping

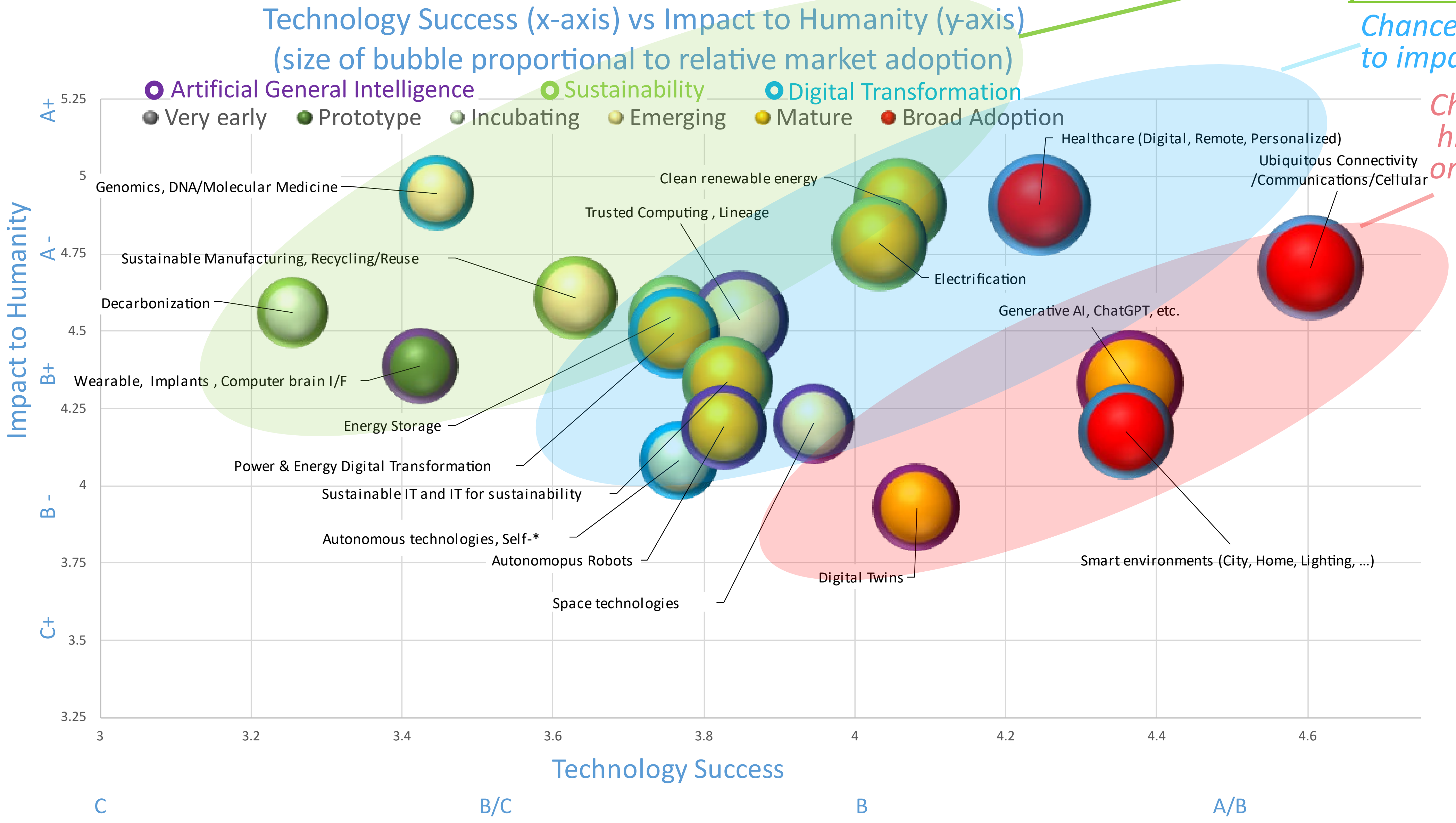


Technology Success (x-axis) vs Impact to Humanity (y-axis)
(size of bubble proportional to relative market adoption)

These are averaged assessments of 48 members of committee

IEEE FUTURE DIRECTIONS

Insights



Technology Success (x-axis) vs Impact to Humanity (y-axis)
(size of bubble proportional to relative market adoption)

Impact on humanity higher than chance of tech success ***<u>(worth investing in)</u>***

Chance of success correlates to impact on humanity

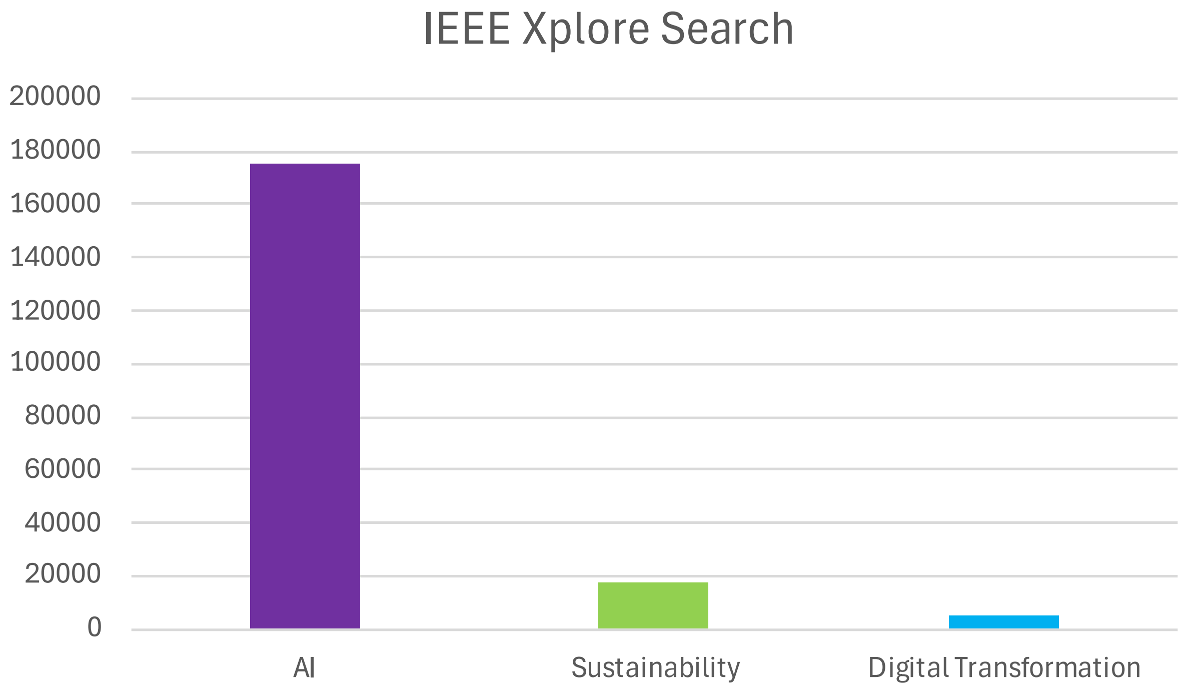
Chance of tech success higher than impact on humanity

Direction of Individual Skills Evolution

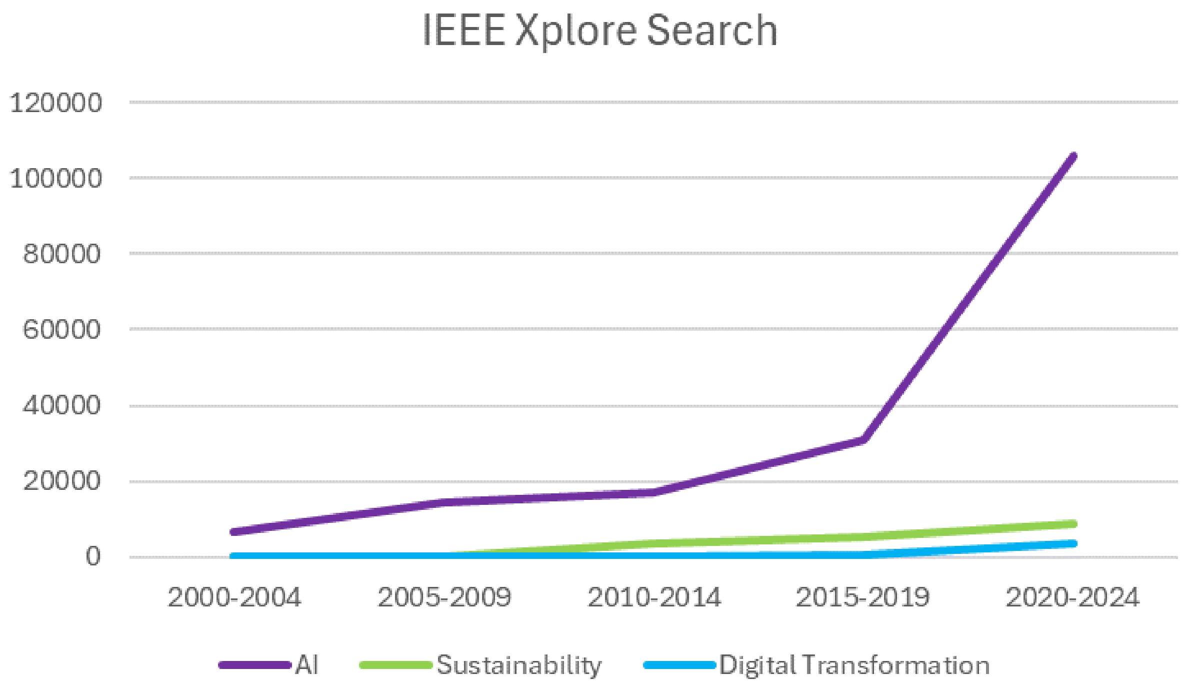
Skills			**Trending**
Digital Transformation	**Sustainability**	**AGI**	
Supervision of automation	Multi-objective optimizations	AI Programmers	↑
Analytics	Measure precursor to manage	Data scientists	↗
Presale, sys integrators	Designers for Sustainability	Solution Architects	→
Maintenance	End-to-end Lifecycle designers	Support	↘
Operators	Sustainability Oversight	System Administrators	↓

This Table was motivated by the table that appeared in K. Bresniker et al., "What Gets You Hired Now Will Not Get You Hired Then," in IT Professional, vol. 26, no. 1, pp. 26-31, Jan.-Feb. 2024. The subset there of, on AGI, was published in the article.

Megatrends vs IEEE Xplore Publications



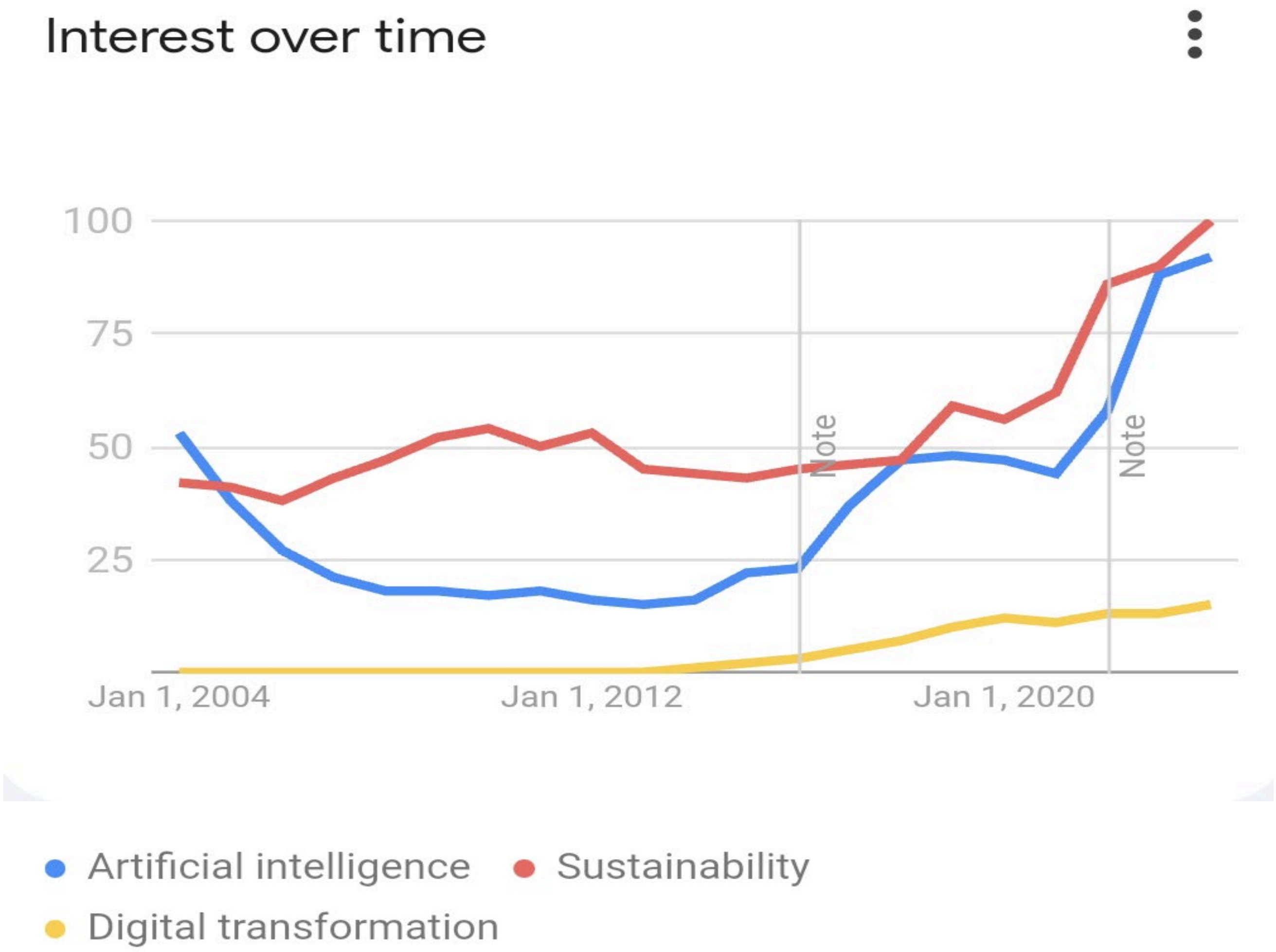
(a) Looked up in January 2024: Overall #documents in IEEE Explore



(b) Looked up in January 2024: #documents in IEEE Explore, growth in each of 5-year segments. Sum of all points are the numbers in (a)

- In publications, AI clearly dominates other two megatrends, this is especially true for the past few years
- We expect this trend will continue in the foreseeable future

Megatrends vs Google Trends

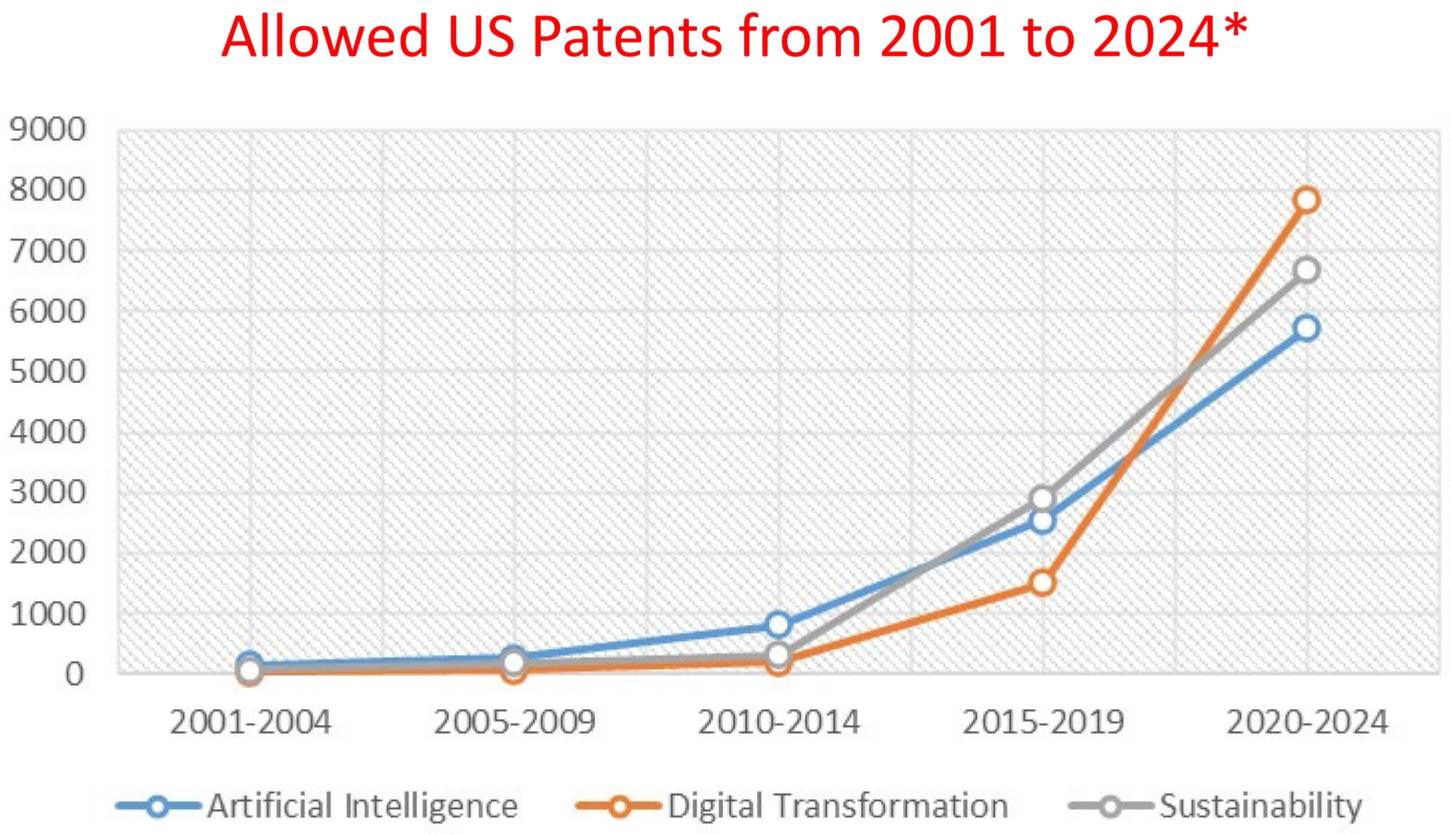


- Surprisingly, sustainability leads among three trends, contrary to AI popularity
- This means that sustainability is firmly on mind of community
- Digital transformation trails substantially which speaks to its maturity

Looked up in January 2024

From Google Trends: *Numbers represent search interest relative to the highest point on the chart for the given region and time. A value of 100 is the peak popularity for the term. A value of 50 means that the term is half as popular. A score of 0 means there was not enough data for this term. (Notes denote dates when improvements to systems were made)*

Megatrends vs US Patents (USPTO)



- AI: there is an upward trend in AI patent filings in recent years, especially from 2015-2019 to 2020-2024.
- Digital Transformation: these patents also show a consistent growth trend with an increase in from 2015-2019 to 2020-2024.
- Sustainability: these patents have witnessed substantial growth from 2010-2014 to 2015-2019.

- Overall, patents trail publications and Google trends. In a way they look backward
- Inherently there is >1.5 year delay from filing to allowing patents
- We expect that patents will catch up in AI domain within ~2 years

*Query conducted in January 2024

General Recommendations

- All three megatrends need to be considered coherently and synergistically
  - A(G)I techniques could be readily applied to sustainable and digitally transformed technologies
  - Sustainability is key aspect of any technology, e.g. AGI requires substantial amounts of processing
  - Digital transformation needs to be continuously modernized taking into account AGI and sustainability
- All three technology megatrends are deeply intertwined with other megatrends and cannot be considered separately
- New Quality of Service (QoS) aspects are being introduced, such as bias, trustworthiness, misinformation, etc.
- Megatrends need to be supported with broad dissemination activity to avoid splitting the society into knowledgeable and left behind.
- One of the challenges is the speed of change being faster than the humans could adapt. This could create fear and aggression. Broad education is critical for technology adoption

Targeted Recommendations

Industry

- Timely productization of near-horizon technologies
- Advance technologies with highest return on investment
- Take responsibility for green technologies
- Make realistic goals and achievable pledges
- Work with academia to educate workforce
- Offer advices to governments how to regulate technology

Government

- Early regulation of technologies that cause concern
- Enforce governance and lineage of data source for training
- Foster research by academia and non-for-profit organizations
- Institute processes and practices against misinformation
- Socialize the mega trends
- Dissemination information for acceptance and explaining risks

Academia

- Globally train trainers for key megatrends
- Work closely with industry to coherently advance science in support of megatrend technologies
- Achieve breakthroughs in fundamental technologies
- Help industry think outside of the box
- Educate (future) workforce of new (mega)trends
- Disseminate materials for all groups/ages for large acceptance

Professional Organization

- Help develop standards suited for increased speed of tech introduction
- Foster communities and events that will address key research problems
- Introduce processes and practices for addressing ethics
- Develop roadmaps for some key technologies of 3 megatrends
- Introduce education, processes, and practices for addressing ethics
- Work closely with industry to better adjust to their needs

Targeted Recommendations, Cont.

End user

- Get acquainted with AI use
- Set expectations correctly
- Green & planet awareness, every little bit helps
- Entertain remote participation instead of flying
- Adopt new devices and tools (that may consume less energy)
- Align with broader infrastructure

Developer

- Get acquainted with AI tools
- Adopt & practice principles of data lineage and trustworthiness
- Focus on sustainable e2e designs
- Make designs observable, verifiable, aligned with SLOs
- E2E Lifecycle awareness
- Minimize data movement
- Any new architecture should be suitable for digital transformation
- Adopt principles of DevOps

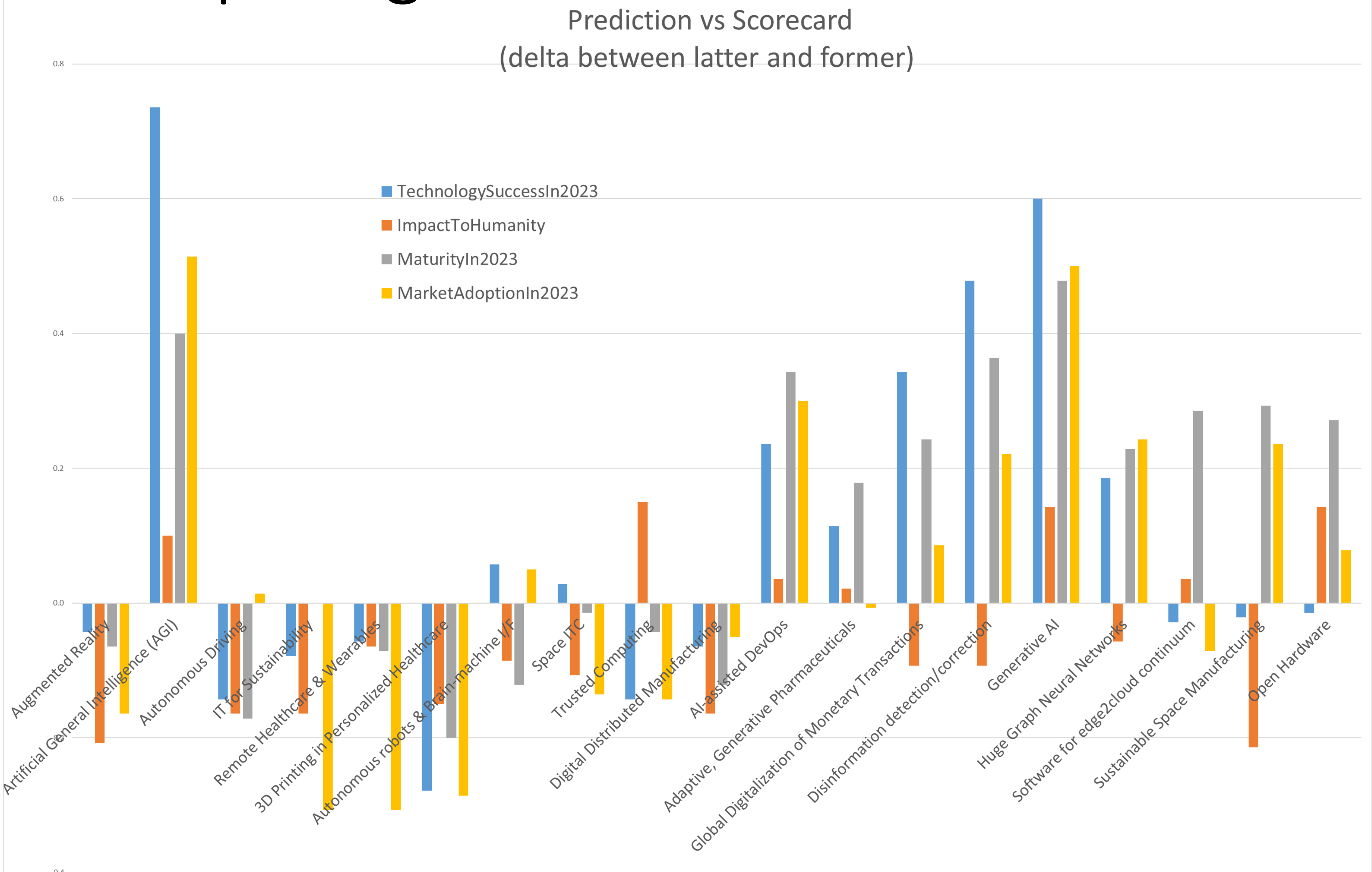
CxO

- Modernize enterprise using AI tools
- Understand AI business and technical risks and opportunities
- Set realistic sustainability expectations
- Carefully align resources to the needs/requirements
- Modernize organization and equipment

Investor

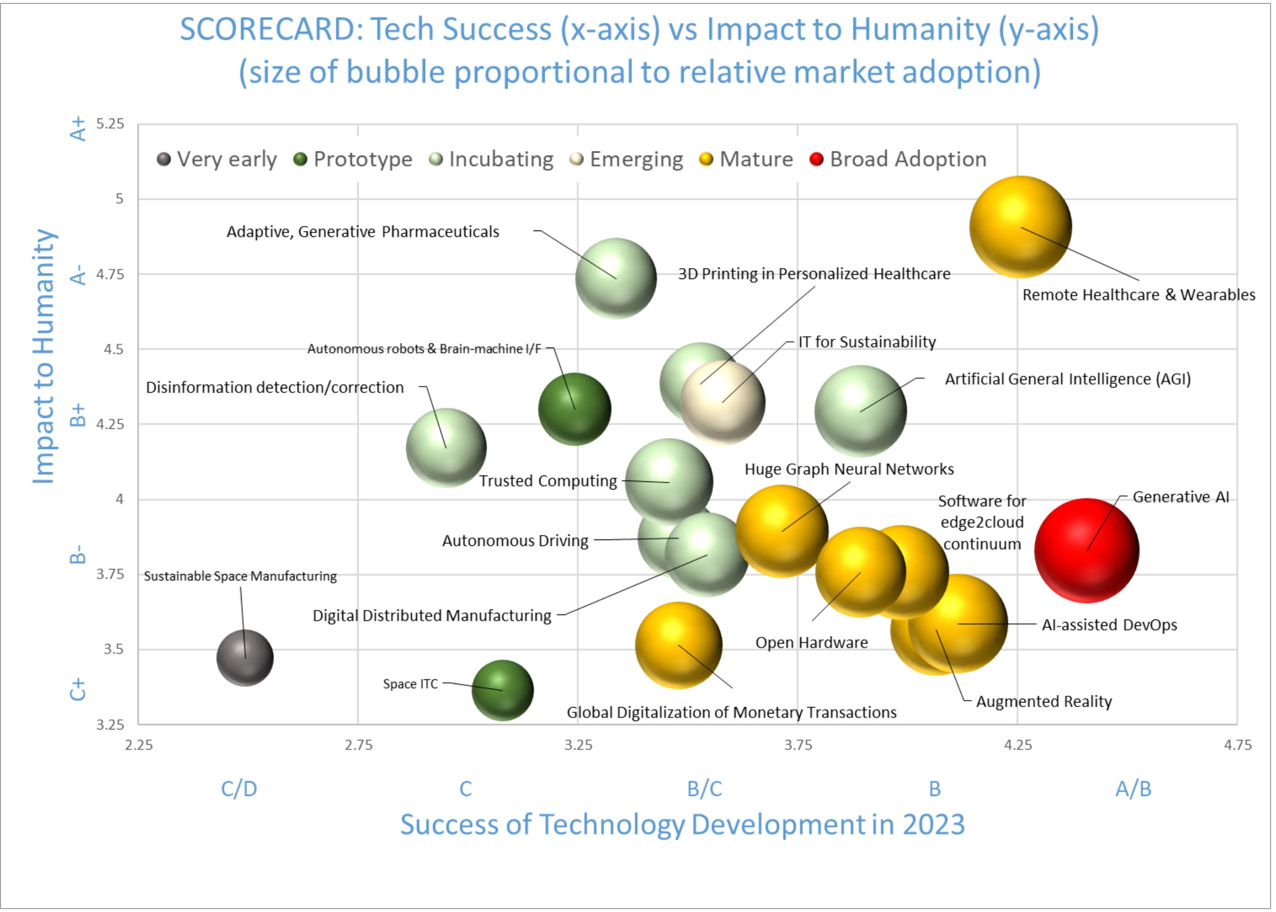
- Invest in balanced tech
- Require coverage of all aspects
- Foster sustainability cross-benefiting green and economy
- Application of AI but not at the expense of sustainability
- Consider new GPUs and new AI accelerators
- Address verticals that have not been digitally transformed

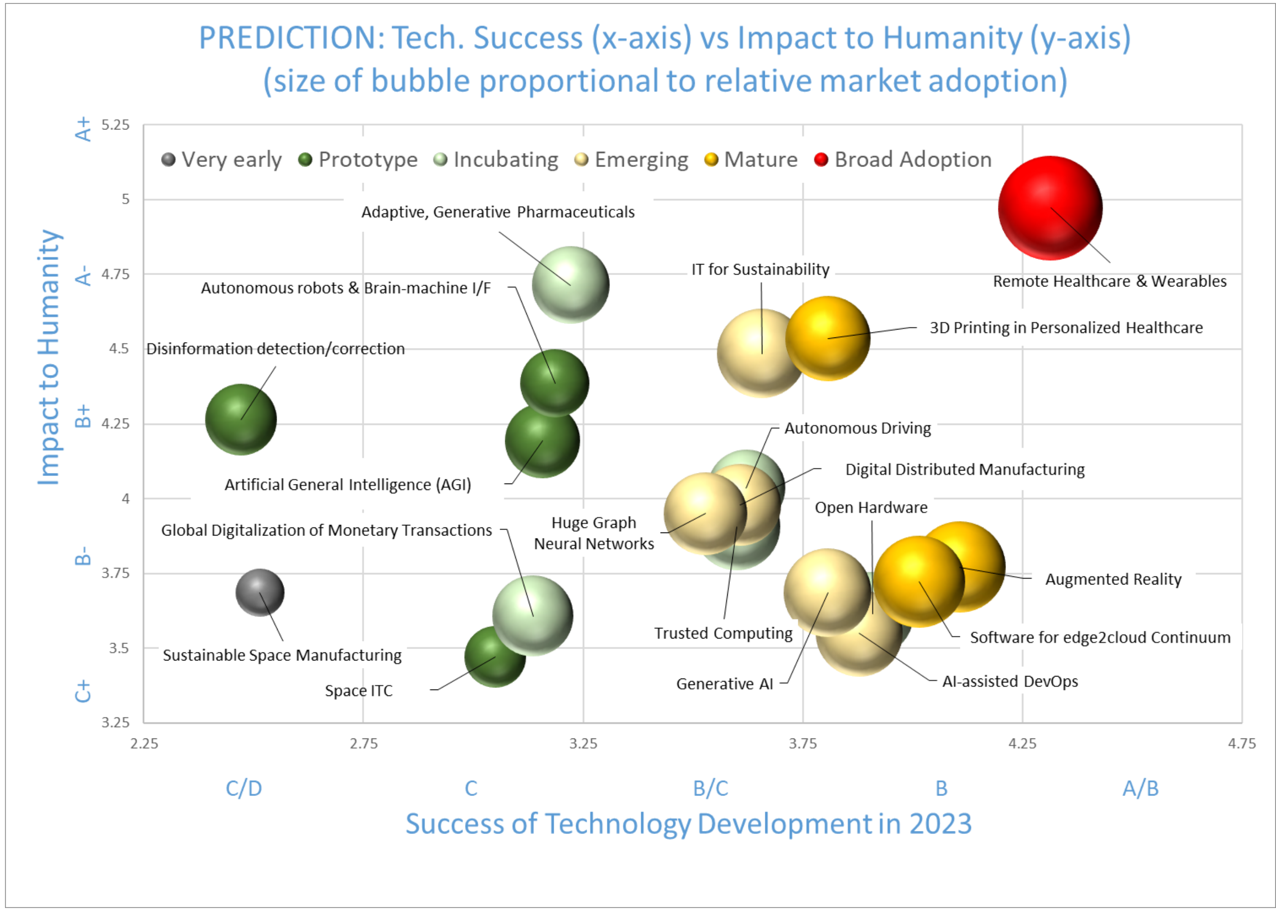
Bars Comparing Scorecard Minus Prediction



Prediction vs Scorecard
(delta between latter and former)

Scorecard grades vs original prediction


SCORECARD: Tech Success (x-axis) vs Impact to Humanity (y-axis)
(size of bubble proportional to relative market adoption)

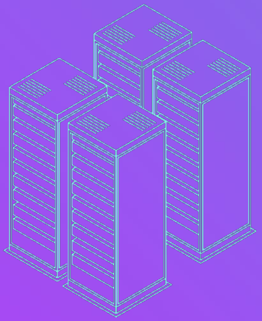

PREDICTION: Tech. Success (x-axis) vs Impact to Humanity (y-axis)
(size of bubble proportional to relative market adoption)

Sustainable data centers

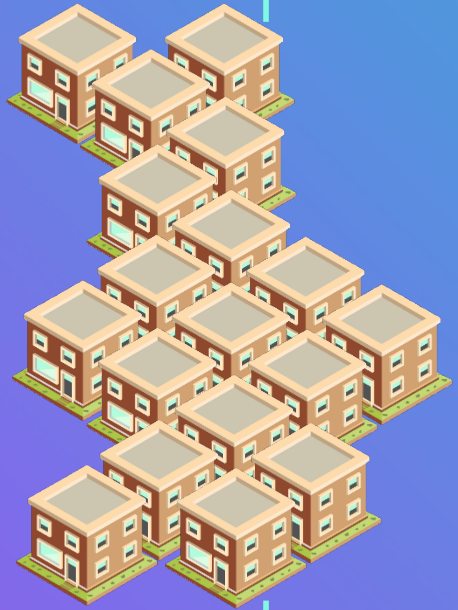
Global electricity consumption could double by 2026

From data centers and AI

Achieving net zero emissions in data centers is daunting, but even more so when faced with the requirements of AI.



Data centers
Artificial intelligence (AI)



By 2026, there could be a

2x

increase from 2022 in data center electricity consumption, to

1,000 TWh

Just **1 TWh** would

- Power 70,000 homes
- Light >1 million homes
- Cool 500,000 homes

…**for one year**

Sources:
IEA report: Electricity 2024: Analysis and forecast to 2026
Duke Energy: Customers surpass 1 terawatt-hour of energy savings

What data center areas can be optimized?

Flexible solutions to determine optimal configurations for each customer

Underutilization
Many operate at low levels of resource utilization, often at only a fraction of capacity

Overprovisioning
Avoiding bottlenecks with more resources can also mean waste when demand fluctuates

Resource imbalance
Balancing workloads across the data center infrastructure can be challenging with changes over time

Environmental impact (carbon & water)

Inefficient cooling
Overcooling or poor airflow can result in excessive energy use and higher operational costs

Slow thermal analysis
Affects planning for the effective placements of IT equipment for optimal sustainability performance

Operational efficiency & reliability

Static vs. Dynamic real-time analysis
Day ahead forecasting is prone to errors, leading to sub-optimal solutions

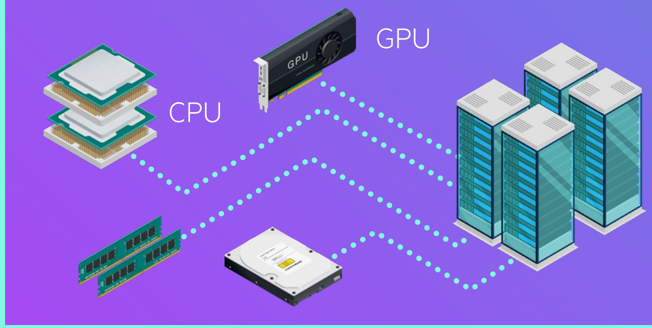
Lack of performance visibility
Makes it challenging to identify inefficiencies and optimize resource allocation

User behavior impact on resource consumption
(using at noon vs midnight, decision for tradeoff)

Hewlett Packard
Labs

The future of sustainable data centers

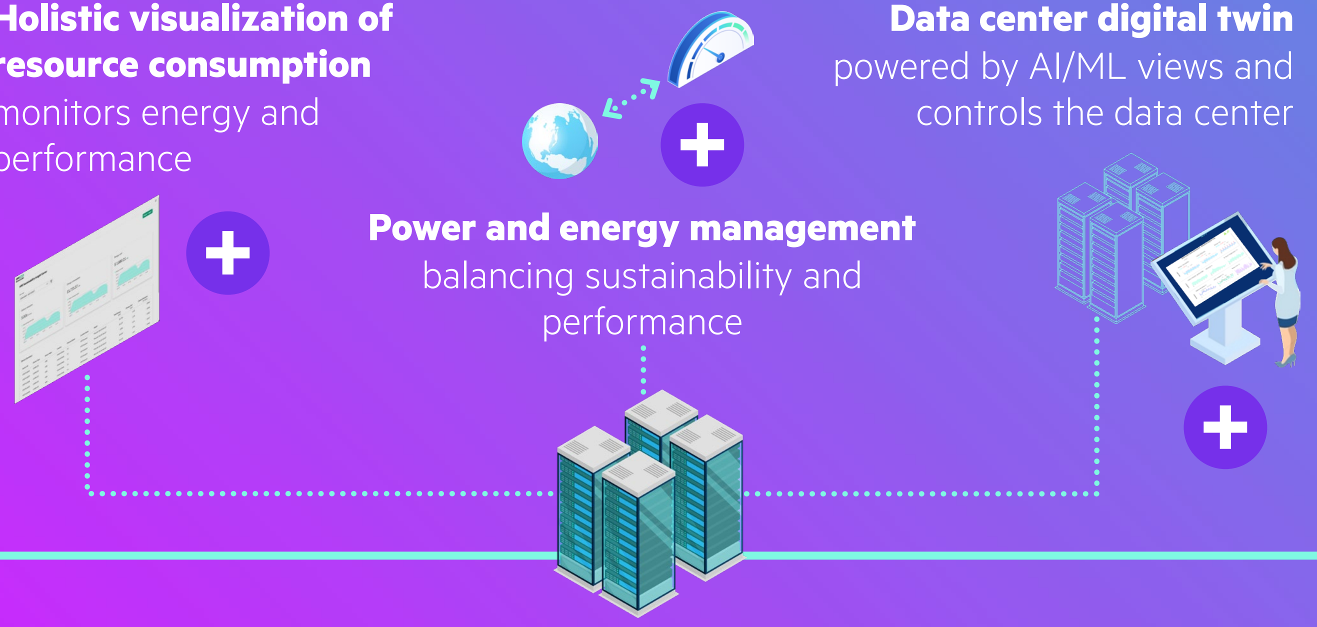
Configuration



Performance-Energy system configuration tool

Configure hardware (virtual + physical) based on workloads

Operations



Holistic visualization of resource consumption monitors energy and performance

Power and energy management balancing sustainability and performance

Data center digital twin powered by AI/ML views and controls the data center

Scaling

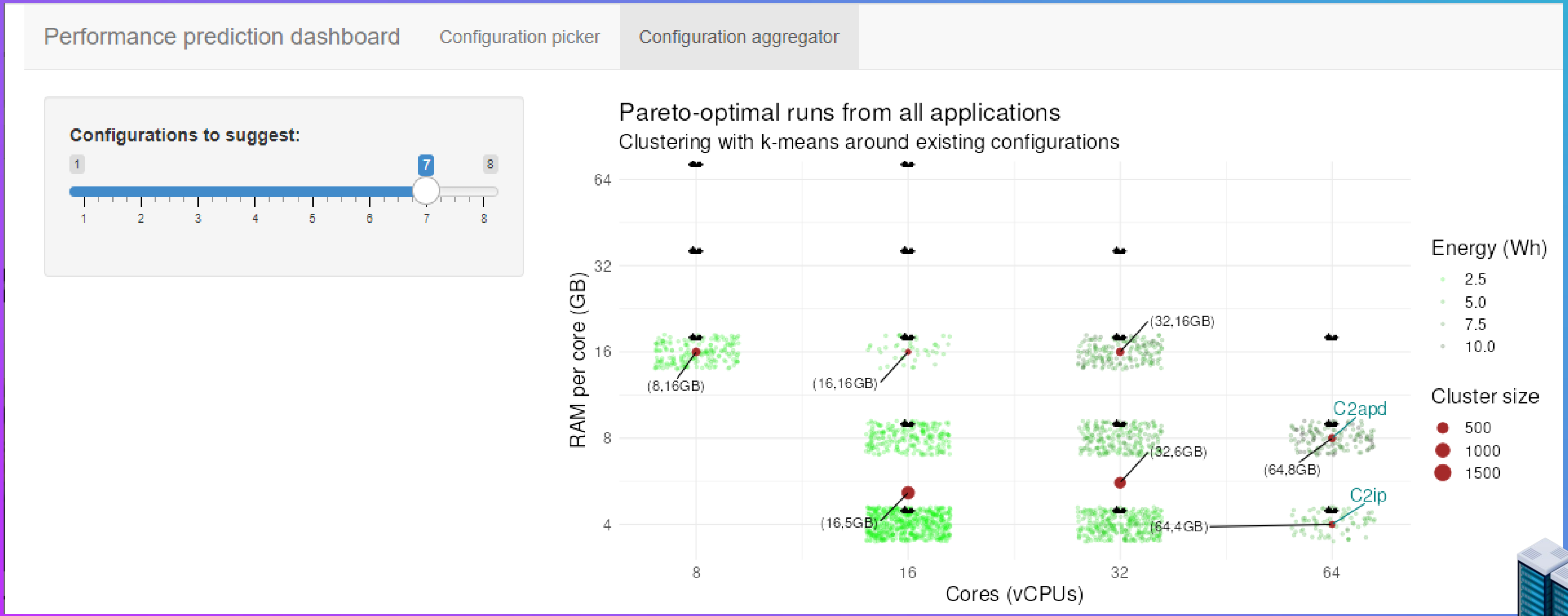
Geo-distributed workload scheduling powered by AI / ML

Hewlett Packard
Labs

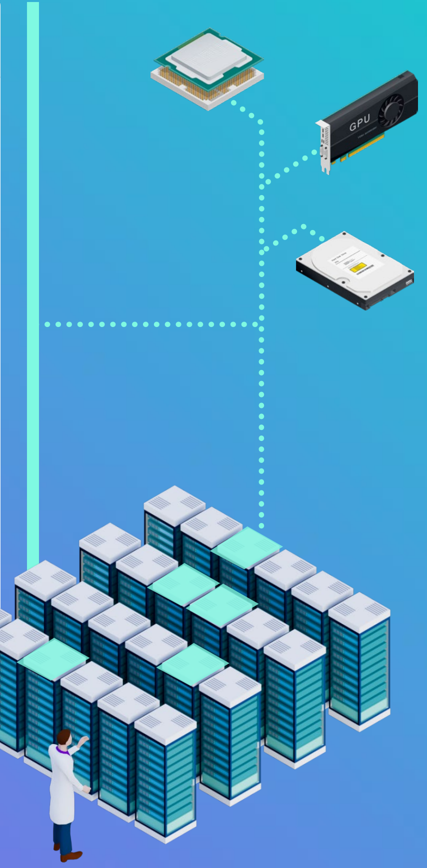
Revolutionizing the future of sustainable data centers

Autoplay ▶

Performance-Energy system configuration tool



Configurations to suggest: 7



Identifies pareto-optimal system configurations
for a set of identified workloads

Hewlett Packard Labs

A Nassereldine, S Diab, M Baydoun, K Leach, M Alt, D Milojicic, I El Hajj, «Predicting the performance-cost trade-off of applications across multiple systems,» Proceedings of CCGrid 2023.

Autoplay ▶

Revolutionizing the future of sustainable data centers

Holistic visualization of resource consumptions

Provides a holistic view of energy profiles:

Carbon footprint

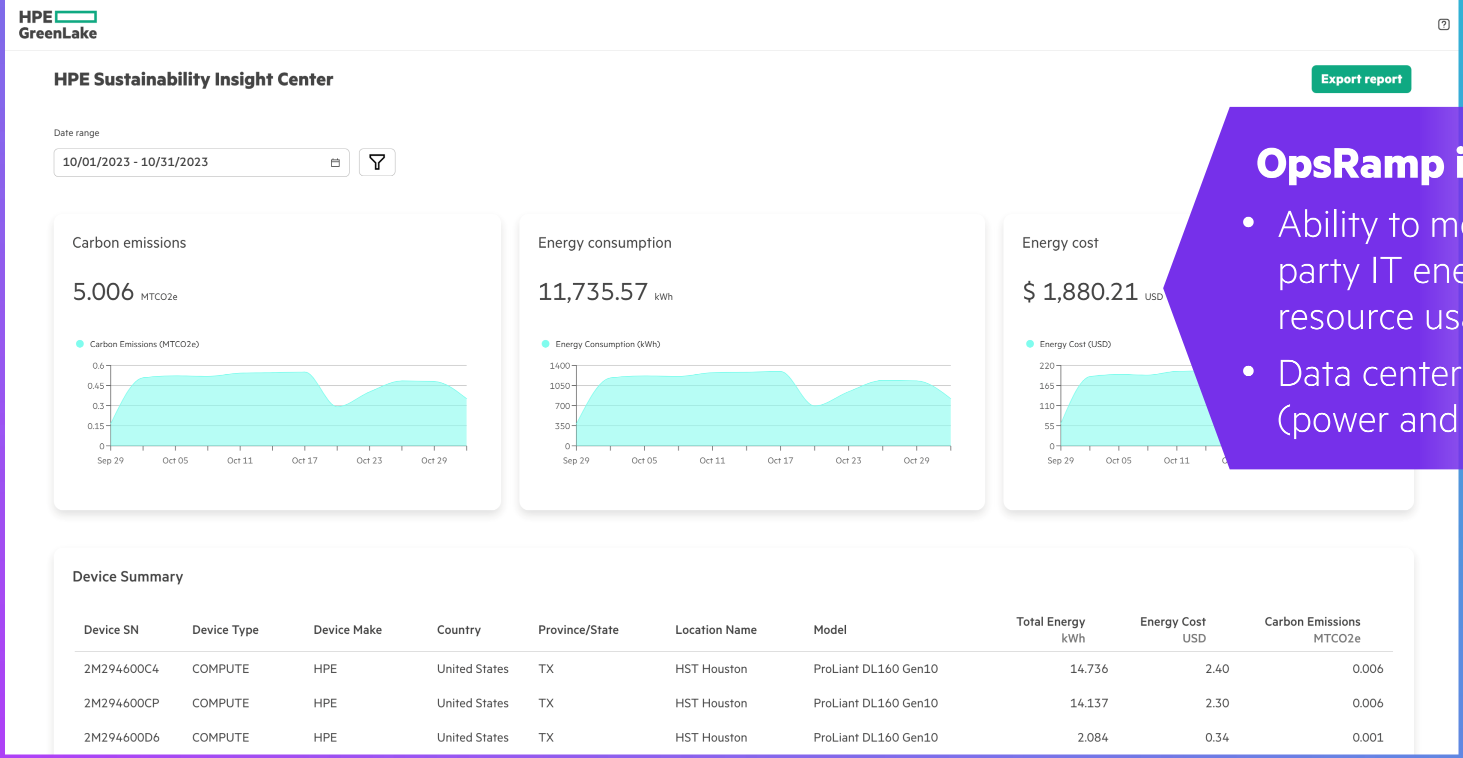
- IT assets based on actual energy usage
- Carbon footprint and energy costs across sites

Devices

Reports energy consumption data from devices

Telemetry

Aggregated across sites (totals and averages)



HPE GreenLake

HPE Sustainability Insight Center

Export report

Date range: 10/01/2023 - 10/31/2023

Carbon emissions: 5.006 MTCO2e

Energy consumption: 11,735.57 kWh

Energy cost: $ 1,880.21 USD

Device Summary

Device SN	Device Type	Device Make	Country	Province/State	Location Name	Model	Total Energy kWh	Energy Cost USD	Carbon Emissions MTCO2e
2M294600C4	COMPUTE	HPE	United States	TX	HST Houston	ProLiant DL160 Gen10	14.736	2.40	0.006
2M294600CP	COMPUTE	HPE	United States	TX	HST Houston	ProLiant DL160 Gen10	14.137	2.30	0.006
2M294600D6	COMPUTE	HPE	United States	TX	HST Houston	ProLiant DL160 Gen10	2.084	0.34	0.001

OpsRamp integration

- Ability to monitor third party IT energy and resource usage
- Data center infrastructure (power and cooling)

Hewlett Packard Labs

HPE Infrastructure and Workload Energy and Emissions Reporting Tool

Main features of the tool

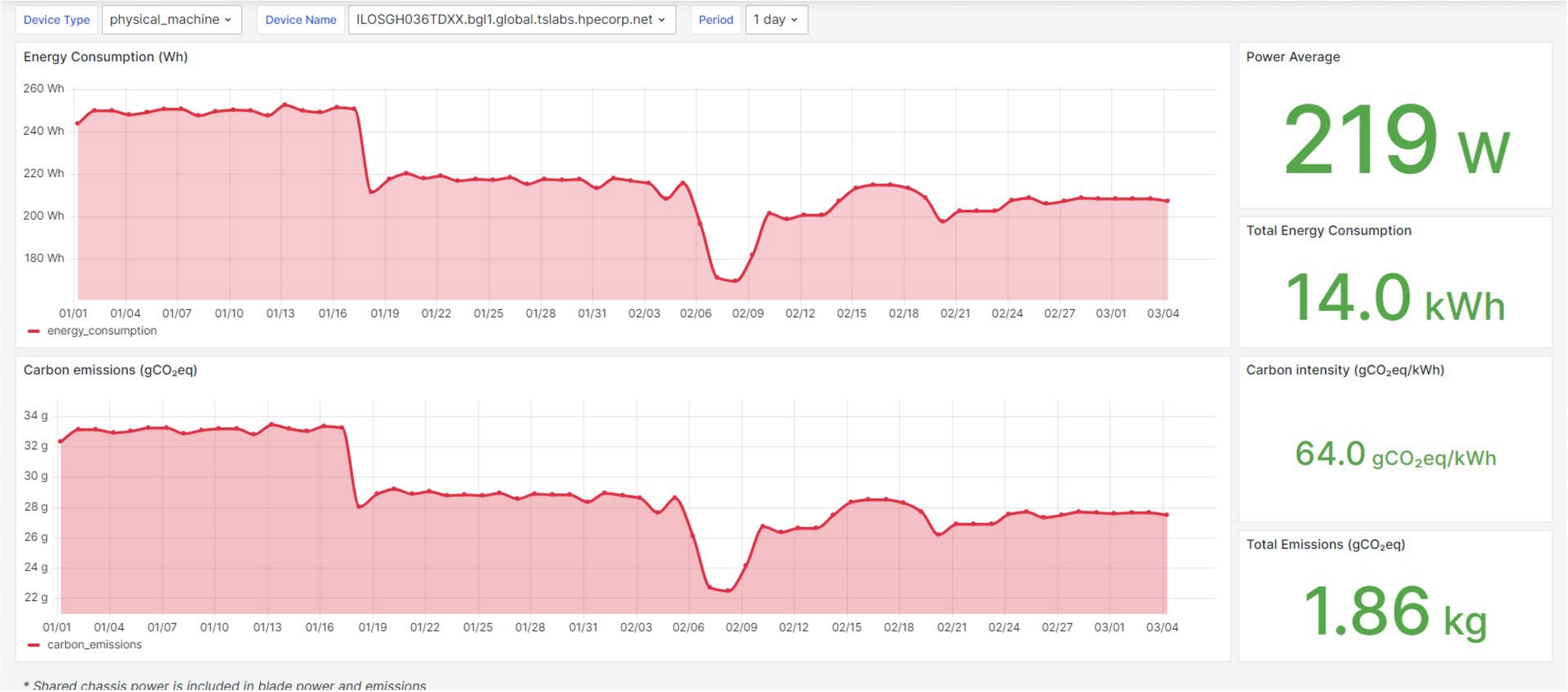
- **Emission Calculator**
  Calculates infrastructure and workload power and emissions using average real power from customer OV-managed physical servers

- **Emissions Estimator**
  Estimates power and emissions based on HPE reference database (HPE and third-party HW)
  For cases when real power data is unavailable

- **BOM Scenario Analysis**
  Shows power and emissions for hypothetical HW scenarios based on location and reference data

Offered by HPE Services as part of Sustainability Services (also known as Greenbird)

Workload emissions dashboard | Infrastructure time series



Device Type: physical_machine

Device Name: ILOSGH036TDXX.bgl1.global.tslabs.hpecorp.net

Period: 1 day

Energy Consumption (Wh)

Power Average

219 W

Total Energy Consumption

14.0 kWh

Carbon emissions (gCO_2eq)

Carbon intensity (gCO_2eq/kWh)

64.0 gCO_2eq/kWh

Total Emissions (gCO_2eq)

1.86 kg

Shared chassis power is included in blade power and emissions

Autoplay ▶

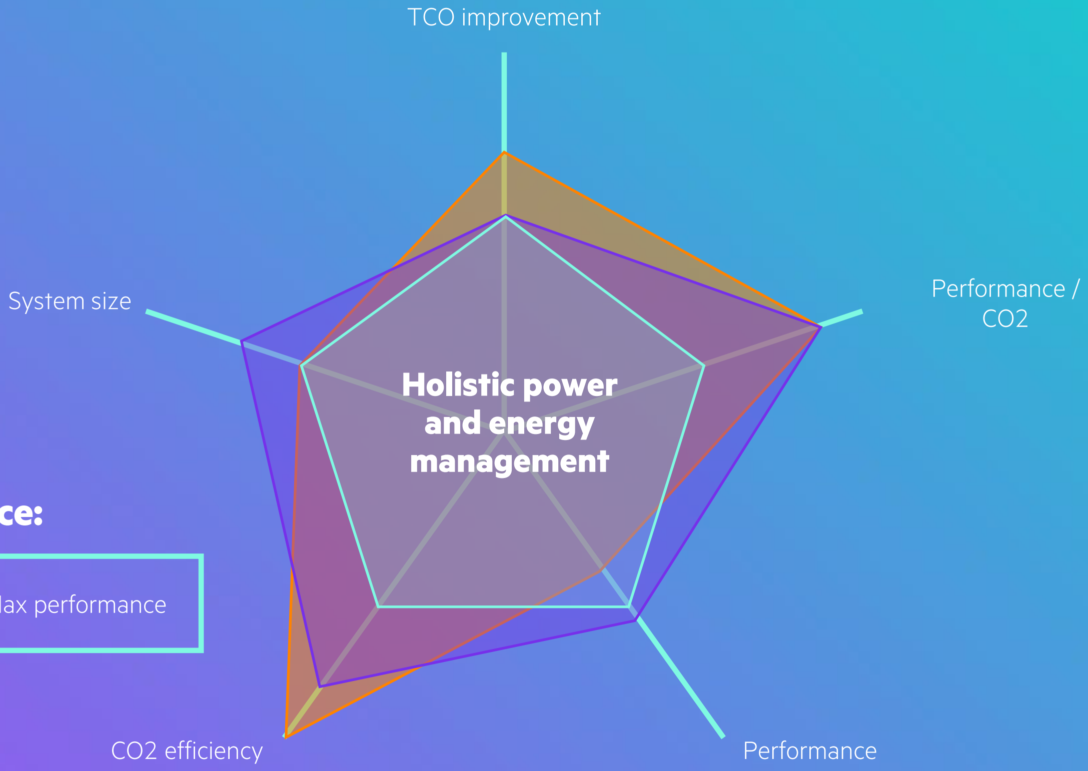
Revolutionizing the future of sustainable data centers

Power & energy management

A pathway to sustainable supercomputing

- Node to system granularity
- Continuous application optimization balancing sustainability and performance
- Accommodates reduction in energy supply while minimizing impact on performance
- Up to 17% energy savings with 6% performance loss for AI workloads

Customer defined preference:



Hewlett Packard Labs

Sustainable Data Center Modernization through Digital Twins

https://www.youtube.com/watch?v=_Js6wXt7tYg

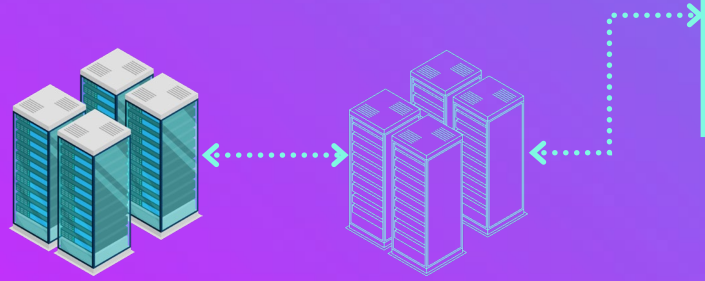
cādence®

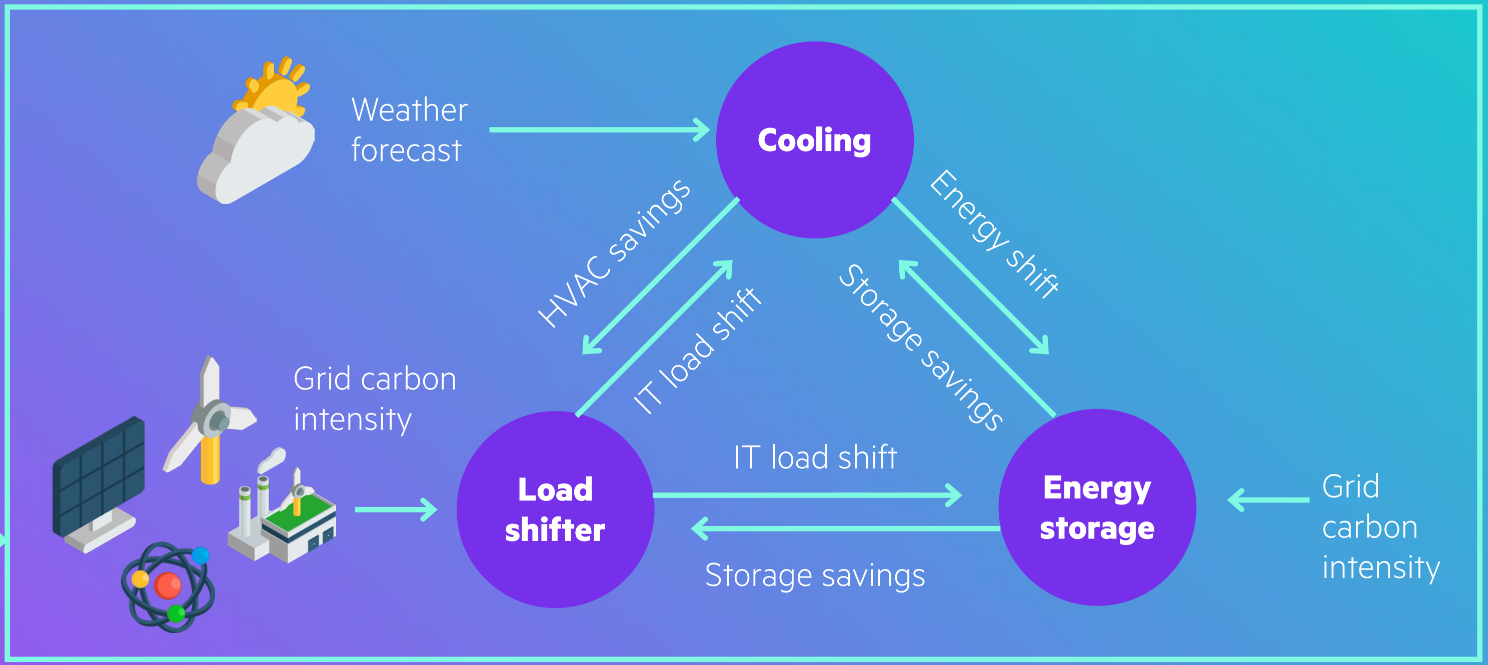
Autoplay ▶

Revolutionizing the future of sustainable data centers

Data Center digital twin

AI with Digital Twins control multiple aspects of the data center in real-time and resolve internal and external dependencies for cooling, load shifting, and battery agents





Sustainable data centers with

- Lower carbon emissions
- Lower energy consumption
- Lower energy cost

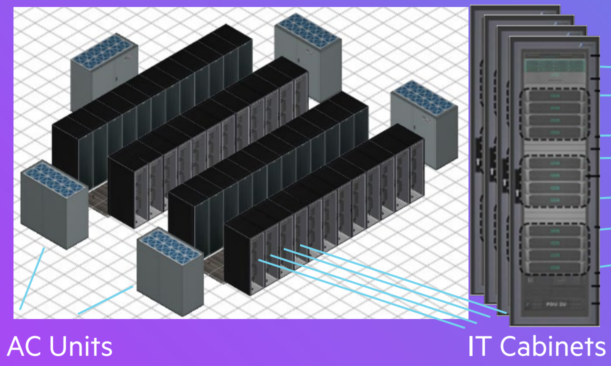
lead to...

Paradigm shift in real-time holistic data center optimization

- Cooling and IT power
- Smart schedule and flexible loads
- Leverage battery storage

Accelerated ML Surrogate Modeling for Cooling Related Analytics

Data Center Configuration

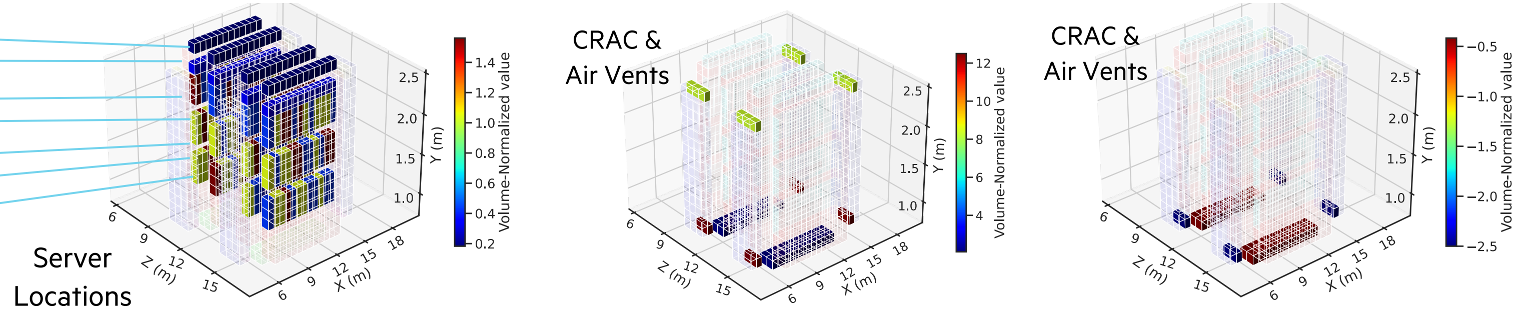


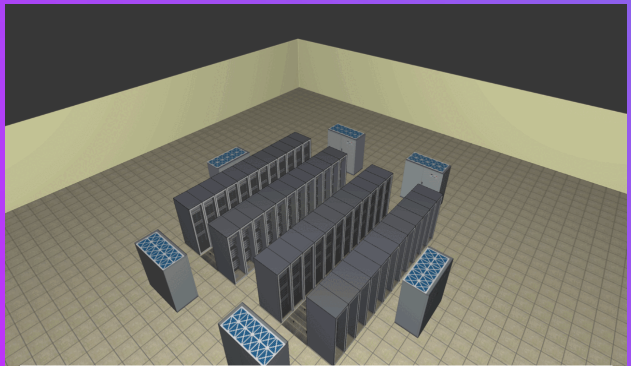
3D CNN Surrogate of CFD

Channel #1: Heat/Power

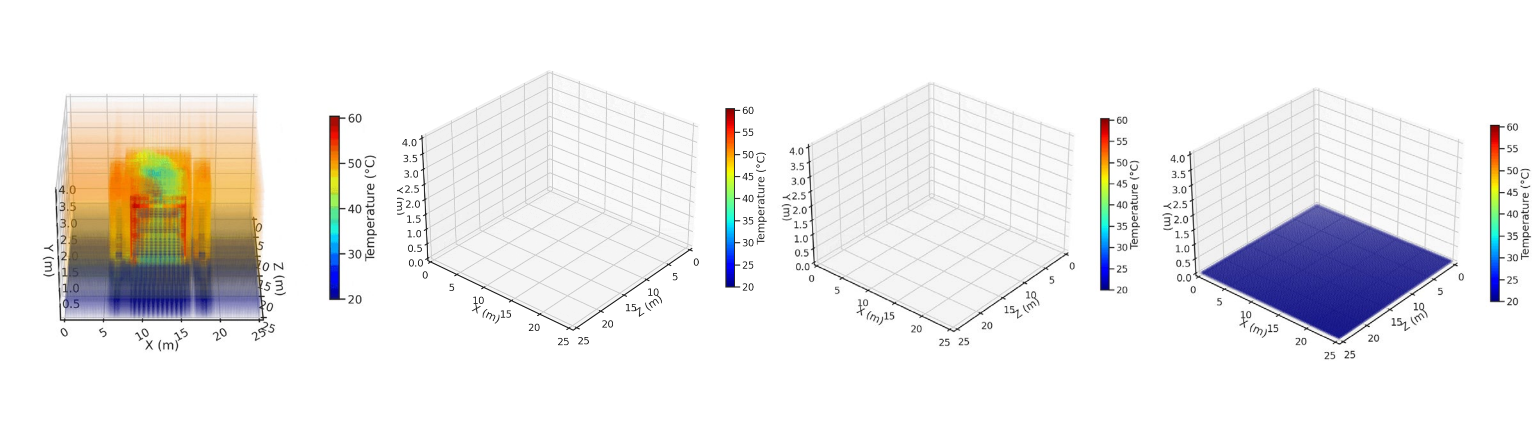
Channel #2: Fan Speed

Channel #3: HVAC Set Point





Data Center Heat Flow

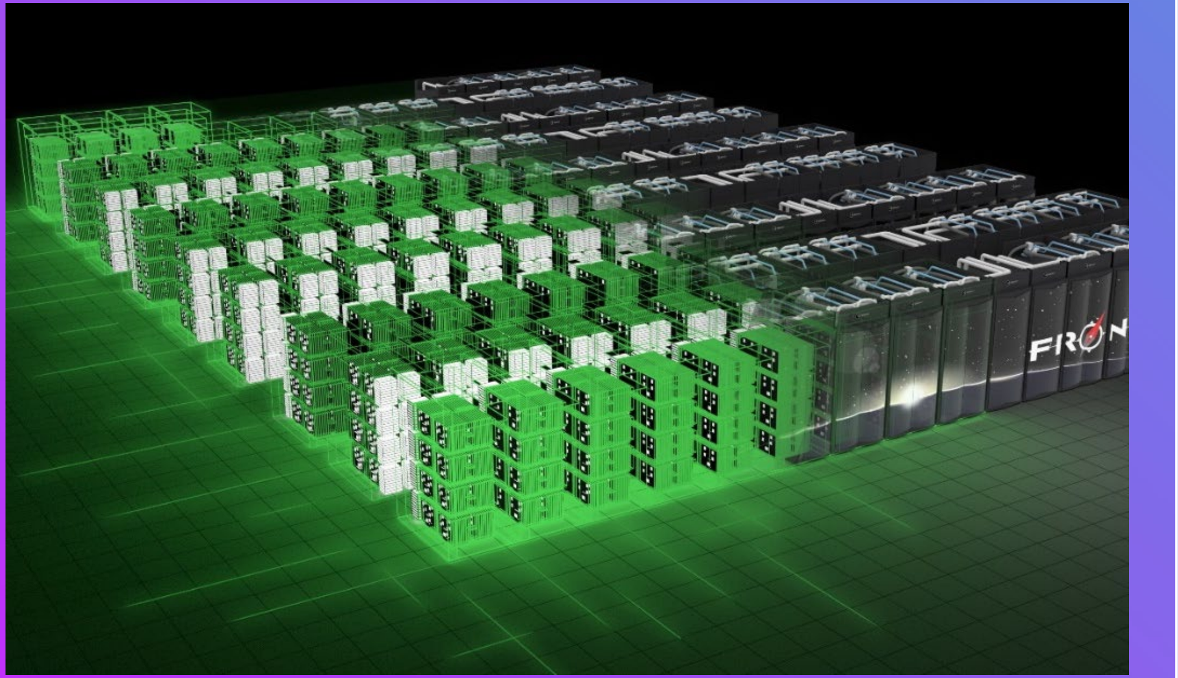


- This will help finer and more effective cooling control of Data Centers saving energy and boost sustainability
- This will help in the design of data centers for the most effective IT cabinet and cooling component layout

S. Sarkar, A. Naug, R. L. Gutierrez, A. Guillen, V. Gundecha, A. Ramesh Babu, and C. Bash, "Real-time Carbon Footprint Minimization in Sustainable Data Centers with Reinforcement Learning," NeurIPS 2023 Workshop on Tackling Climate Change with Machine Learning. [Best Paper Award for ML Innovation] https://www.climatechange.ai/papers/neurips2023/28

EXADIGIT Project:
Digital Twin consortium for supercomputing

Building an open-source community for comprehensive modeling of supercomputers



Collaboration with Oak Ridge National Laboratories

Finland
CSC – IT Center for Science

France
INES

Sweden
KTH Royal Institute of Technology

Czech Republic
VSB Technical University of Ostrava / IT Innovations National Supercomputing Center

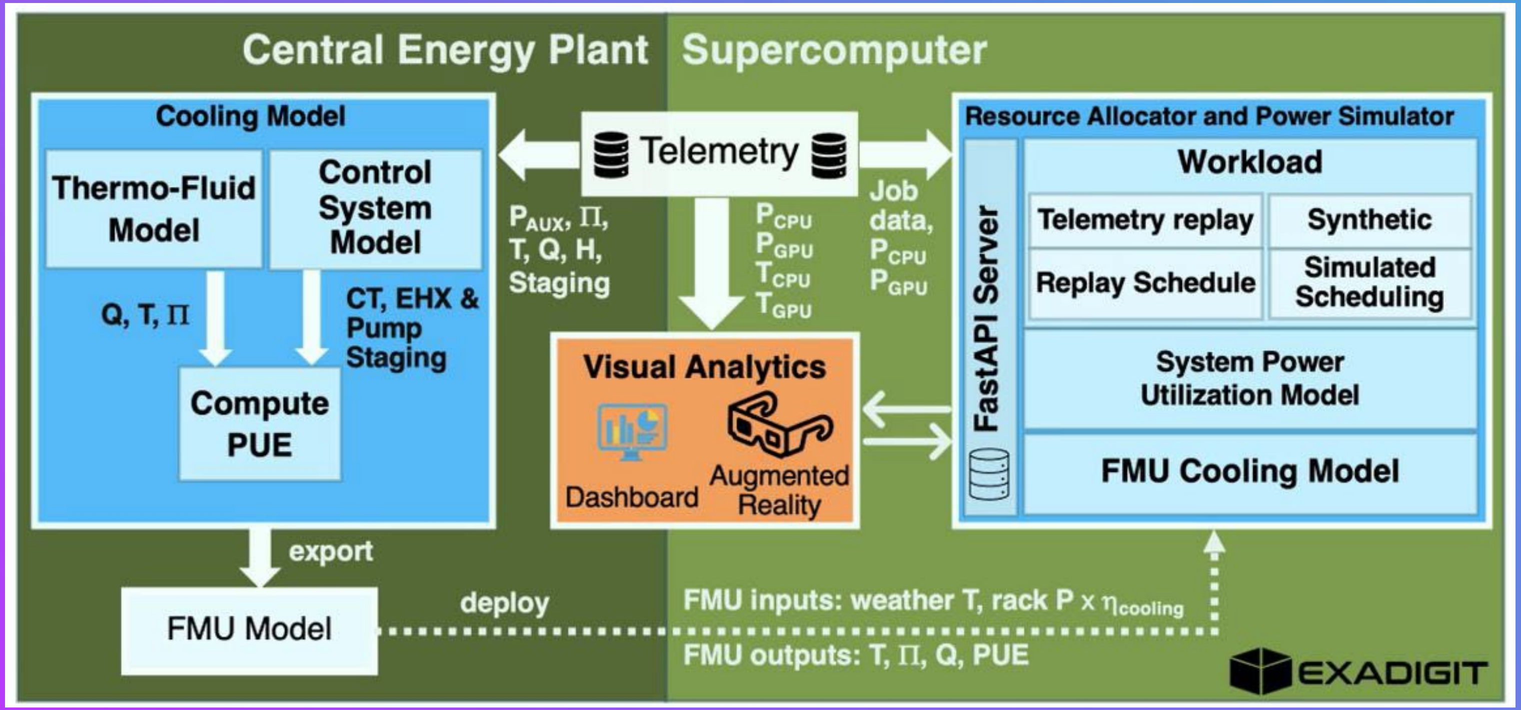
United Kingdom
EPCC

Australia
Pawsey

Germany
Jülich Forschungcentrum

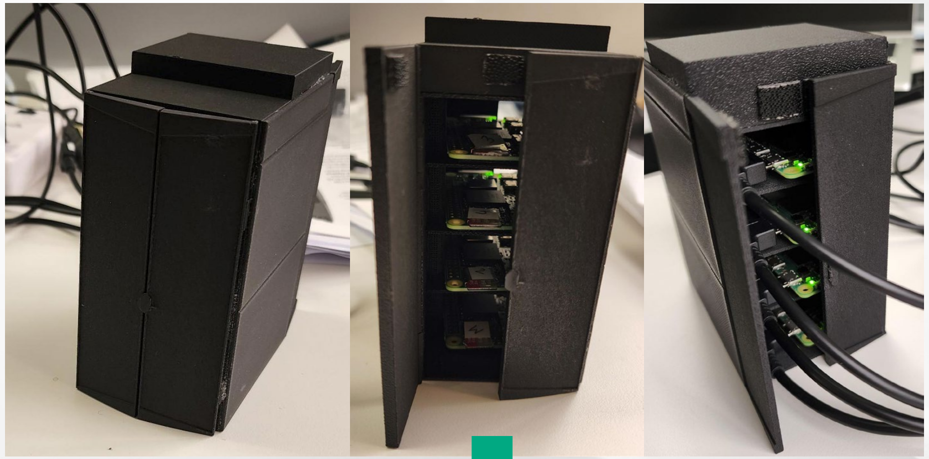
Industry partners
Hewlett Packard Enterprise
NVIDIA

ExaDigiT Architecture (Evolving)



W. Brewer, M. Maiterth, V. Kumar, R. Wojda, S. Bouknight, J. Hines, W. Shin, J. Webb, S. Greenwood, W. Williams, D. Grant, and F. Wang, “A Digital Twin Framework for Liquid-cooled Supercomputers as Demonstrated at Exascale”, in Proceedings of the International Conference for High Performance Computing, Networking, Storage, and Analysis (SC’24), 2024.

Mini-Frontier Digital Twins: Demo At Discover

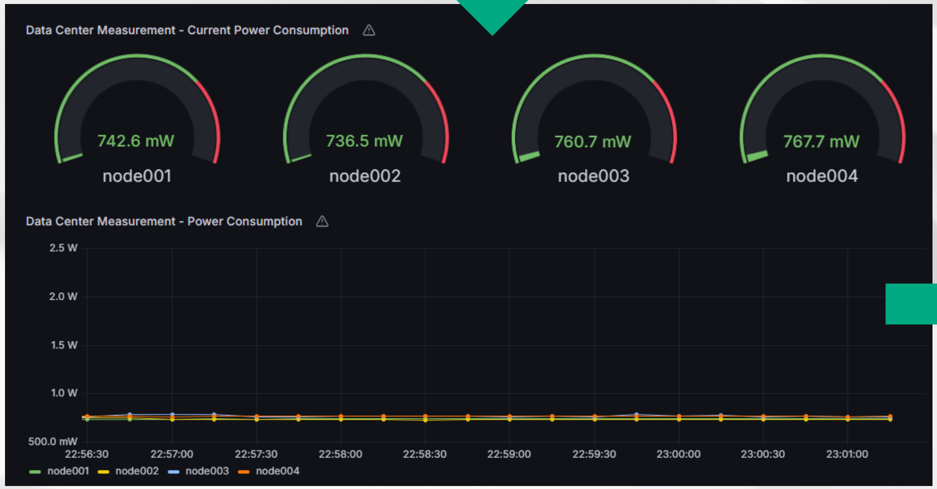


Summary:

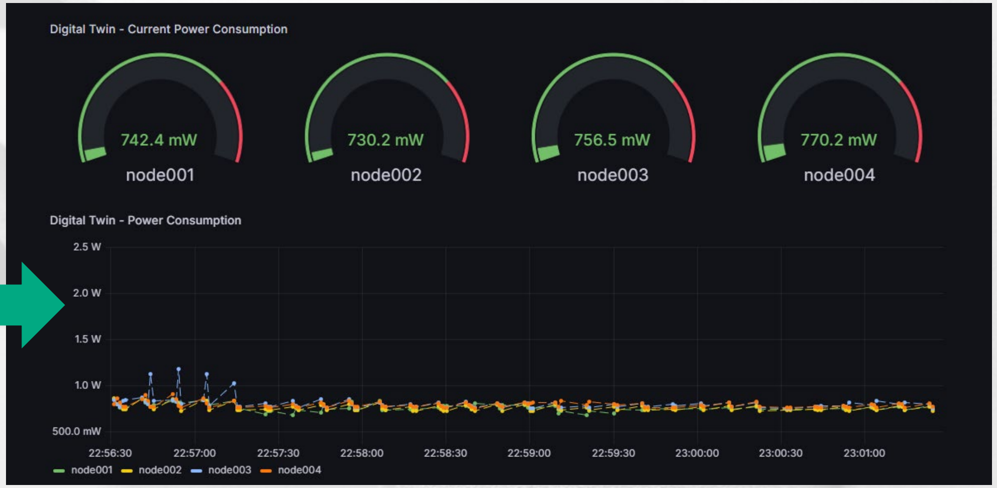
- We fitted a 3D-printed mini-Frontier cabinet with 4 Raspberry Pies and ran mini-HPC workloads (as seen on the left)
- As part of a larger project, we built a Digital Twin monitoring system that provides a dashboard of metrics such as power consumption, CPU, and memory usage, and predicts future loads in the next few minutes

Highlights:

- Over 100 people stopped by our booth
- Mini Frontier was a big attraction



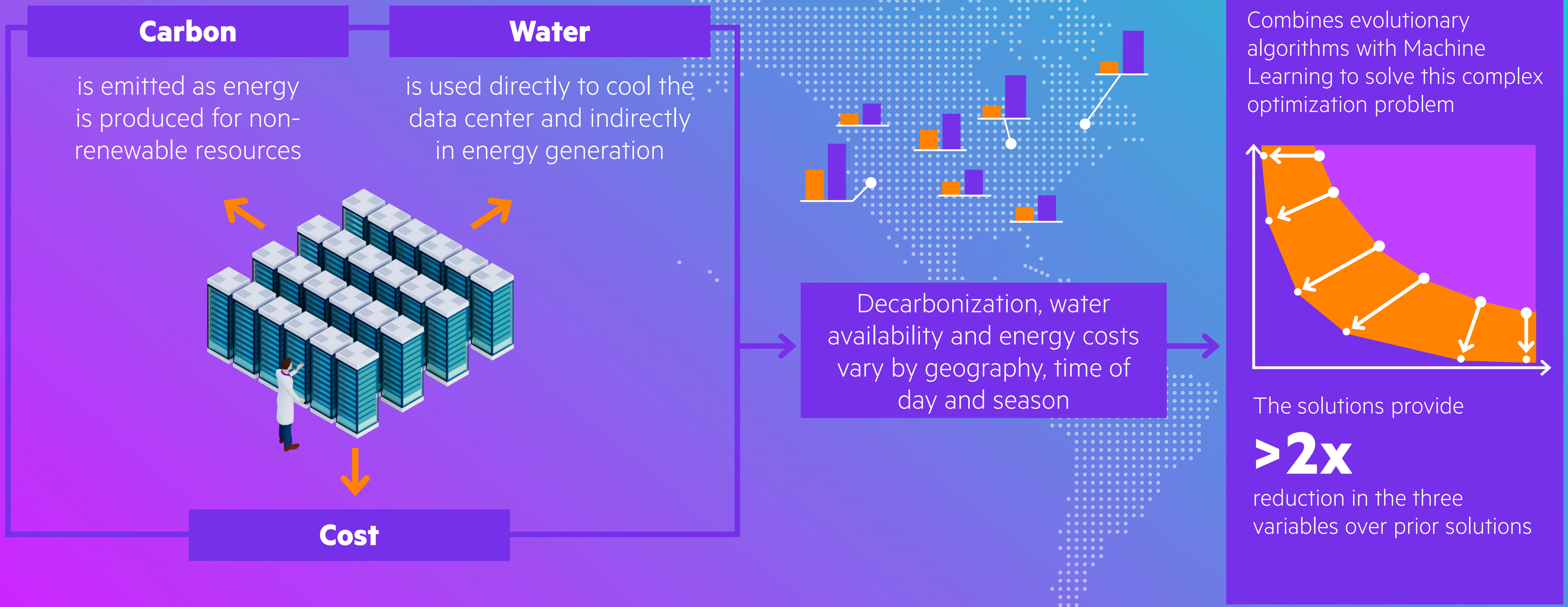
Power Consumption Real-Time Monitoring Dashboard



Digital Twin monitoring Dashboard

Revolutionizing the future of sustainable data centers

Geo-distributed workload scheduling



Carbon

is emitted as energy is produced for non-renewable resources

Water

is used directly to cool the data center and indirectly in energy generation

Cost

Decarbonization, water availability and energy costs vary by geography, time of day and season

Combines evolutionary algorithms with Machine Learning to solve this complex optimization problem

The solutions provide

>2x

reduction in the three variables over prior solutions

Sirui Qi, Dejan Milojicic, Cullen Bash, and Sudeep Pasricha, "SHIELD: Sustainable Hybrid Evolutionary Learning Framework for Carbon, Wastewater, and Energy-Aware Data Center Management," IEEE 14th IGSC, 2023. [Best paper award]

Summary

- At least in the foreseeable future, AI is driving computing in general and HPC in particular
- Sustainability from economical and ecological perspective is critical to deliver this new computing
- We are all in the same boat (end users, developers, providers, integrators, …), we all need to act
- Holistic, end-to-end, perspective is important

Thank you

Dejan Milojicic, HPE Fellow, Systems Architecture Lab
dejan.milojicic@hpe.com
Cullen Bash, Vice President & Director, Systems Architecture Lab
cullen.bash@hpe.com

© 2024 Hewlett Packard Enterprise Development LP