Build Generative AI operations at scale

Abhishek Agarwal

05-17-2024 | ProductCamp RTP

About Me

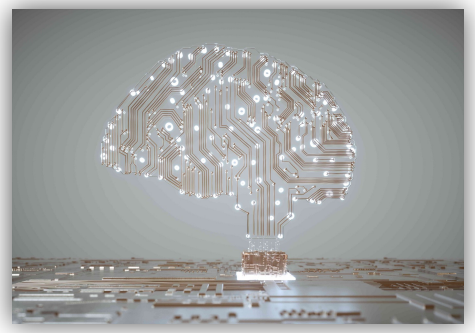
Abhishek Agarwal

+ 1 (475) 239-8726
abhishek.agarwal9219@gmail.com
New York,, NY

Strengths

- Leadership
- Innovation
- Collaborate
- Adaptability
- Learning

Focus



Responsible ML/AI Product Development

Education

MBA
Yale SOM

B. Tech (Hons)
IIT Kharagpur

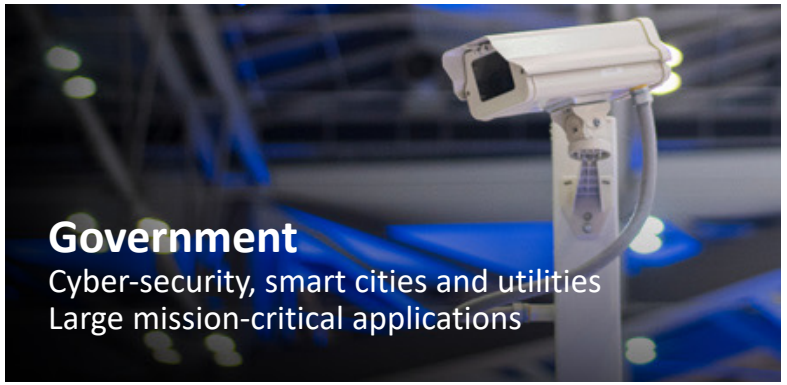
Experience

- HPE
- Amazon
- Unilever

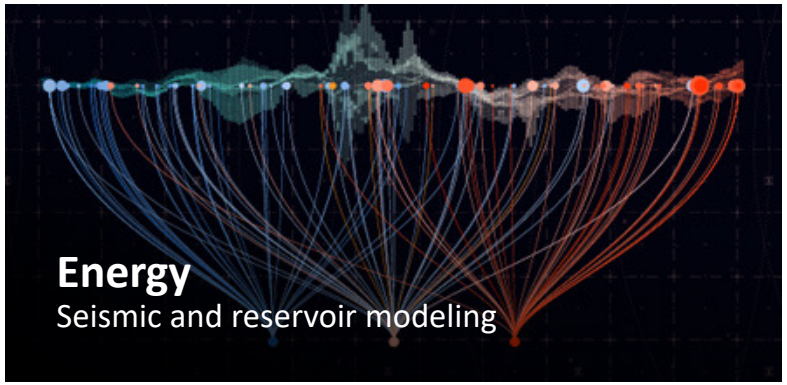
Personal Interest

- Born & raised in India
- Fun fact: Goal of visiting every continent
- Extend access of technology
- Promote sustainable development

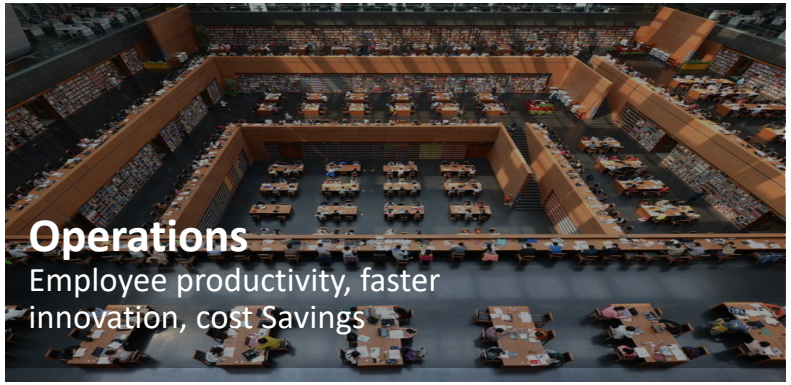
Financial services
Fraud detection, ID verification



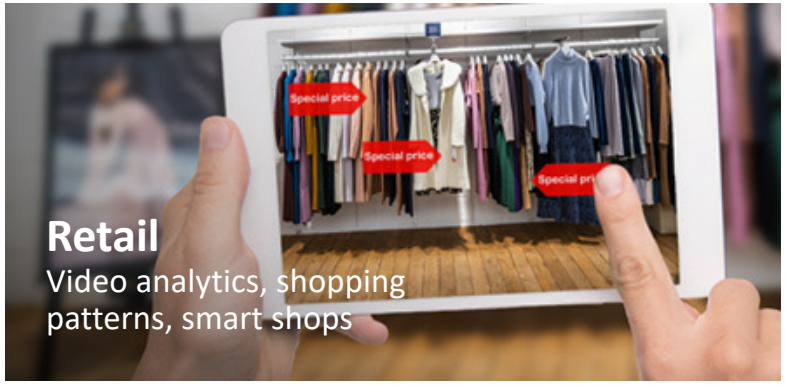
Government
Cyber-security, smart cities and utilities
Large mission-critical applications



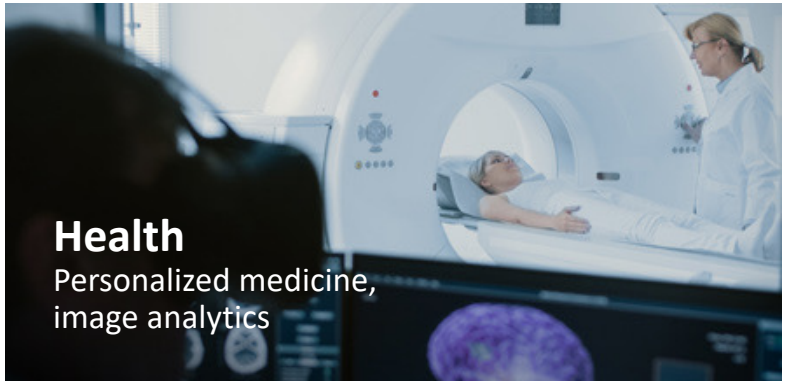
Energy
Seismic and reservoir modeling



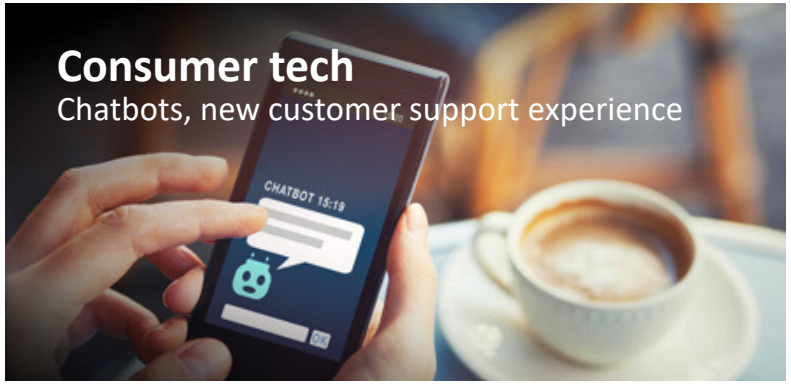
Operations
Employee productivity, faster innovation, cost Savings



Retail
Video analytics, shopping patterns, smart shops



Health
Personalized medicine, image analytics

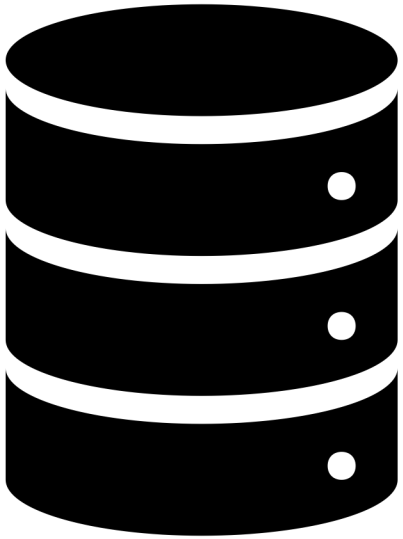


Consumer tech
Chatbots, new customer support experience

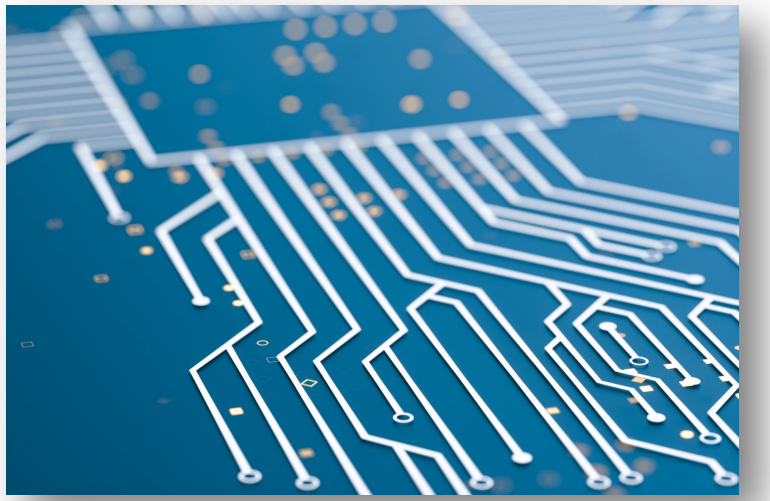
Service providers
Network monitoring, support automation, IT service management

Manufacturing
Quality control, video analytics and prescriptive maintenance

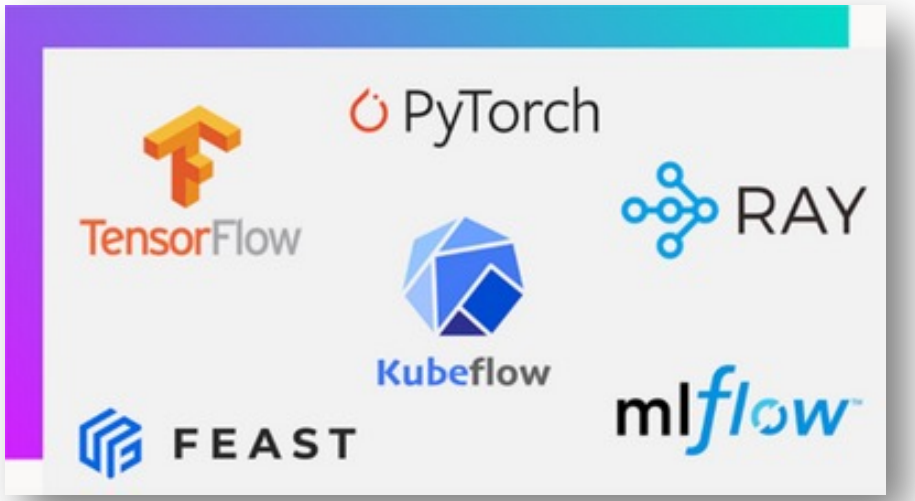
Generative AI Journey



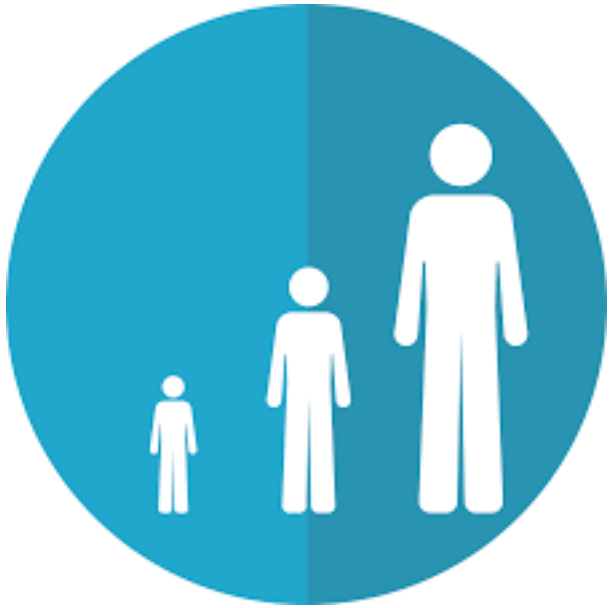
Data

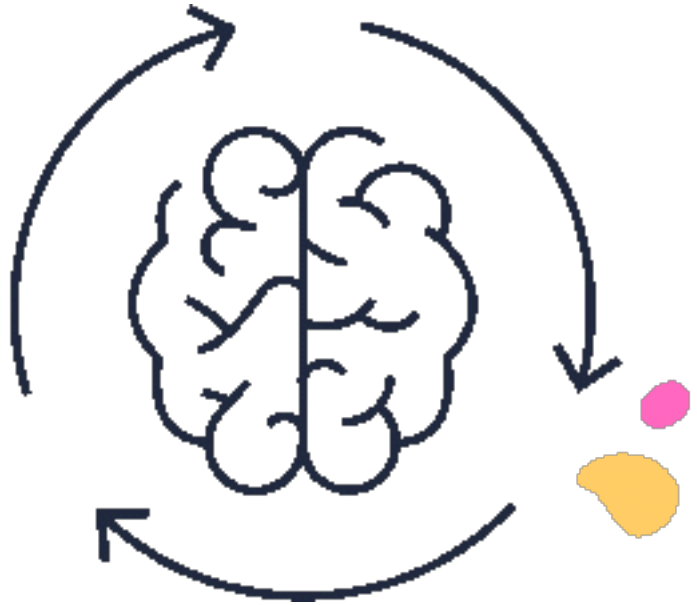


Compute



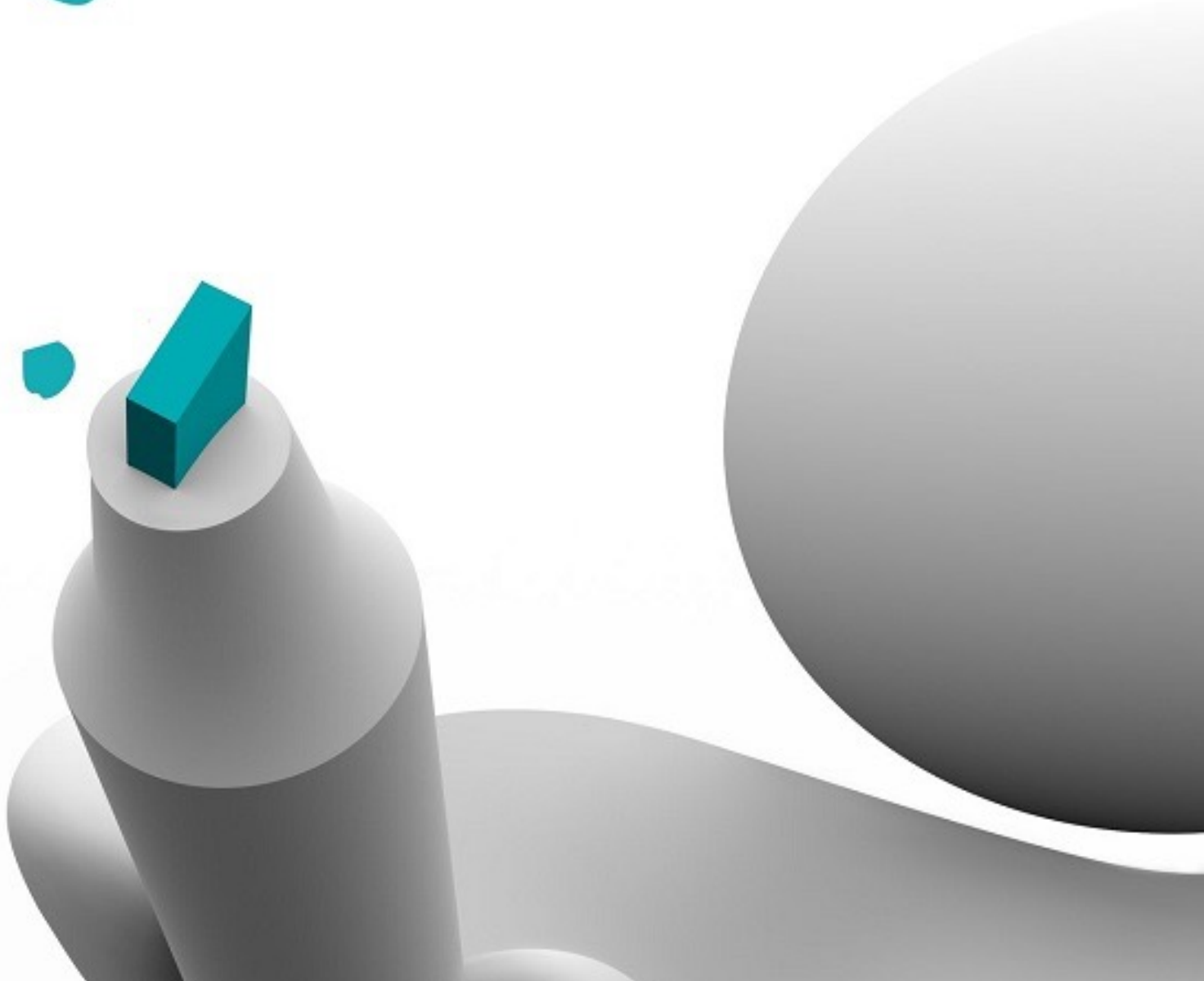
Algorithms





Human Lifetime	Foundational Models
1 Billion Words	Trillions Words

HOW TO...



Generative AI Adoption Maturity Levels

Business Value Generated at every Level

01 Emerging Level

- Build training programs for business and technology teams
- Business experiments with prompts on nonproprietary use cases on market lea open/closed LLM.
- Technology teams help with making data ready for Model consumptions

02 Progressive Level

- An AI Policy is established to ensure compliance..
- Business is able to use a hosted model to solve their proprietary use cases
- Human prompts are enhanced programmatically by systems

03 Growing Level

- Retrieval Augmented Generation based Architecture in place
- Finetuning of Models becomes normal, and ML Operations are operationalized.
- Agents are integrated with Models to integrate LLMs with enterprise applicatior

04 Optimizing Level

- LLMs are smaller and finetuned to run on commodity hardware
- Multi model data preparation tools in place and baked into ML Ops to make dat model ready
- Costs of running Models decreases while adoption in the enterprise increases

Generative AI – Product Managers

Product Development

Large Language Models
Machine Learning
Data
Code Generation
Interfaces & processes
Solving complex problem

Product Features

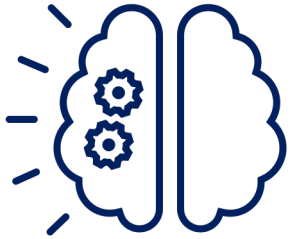
Capabilities
Values (better, faster, cheaper)
Interaction

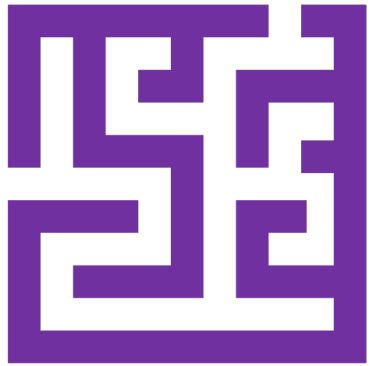
Product Support

Customer Service
FAQs
Self improving

AI – Value Impact

Ethics





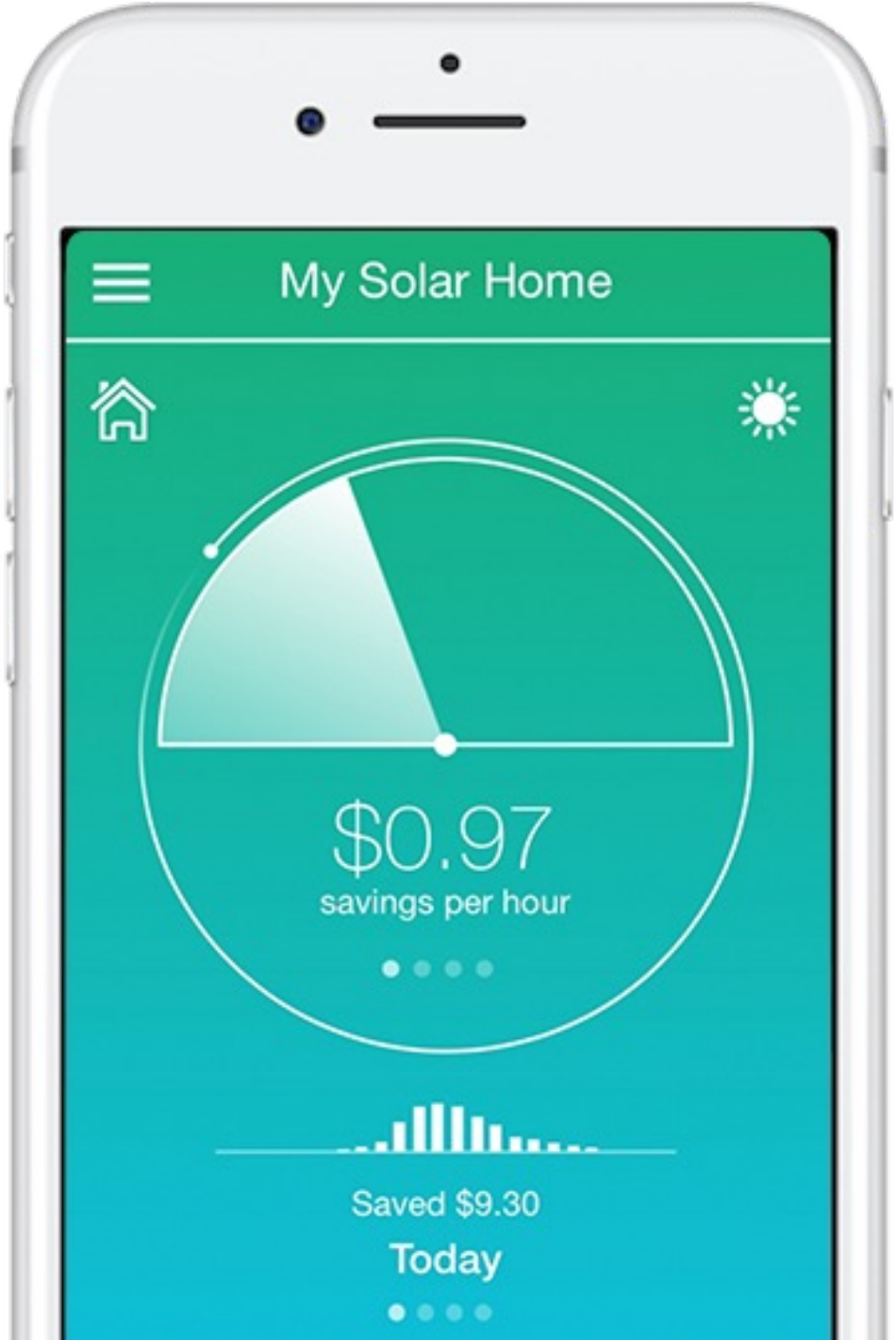
Ethical complexity

Societal Impact

Human centric design

Scenario

An AI designed to optimize energy usage in homes is discovered to potentially compromise user privacy by tracking when residents are home or away. As a product manager, how do you respond?

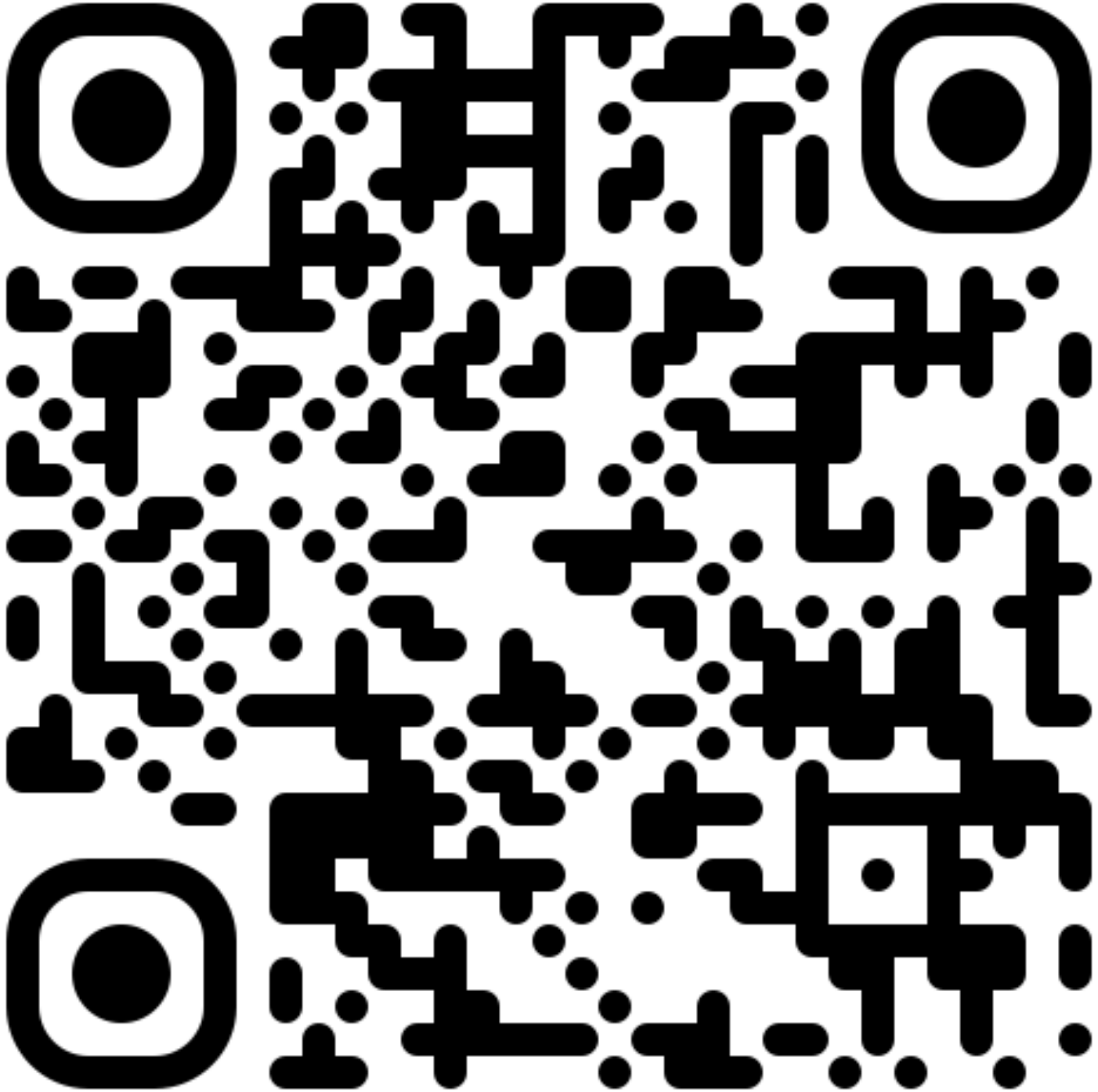


slido

An AI designed to optimize energy usage in homes is discovered to potentially compromise user privacy by tracking when residents are home or away. As a product manager, how do you respond?

ⓘ Start presenting to display the poll results on this slide.

Best Practices

Thank you.
Connect with me!

https://www.linkedin.com/in/abhishekagarwal92/