NSF: Secure and Trustworthy Cyberspace Frontiers

Center for Distributed Confidential Computing (CDCC)

XiaoFeng Wang (Director)

xw7@iu.edu

https://homes.luddy.indiana.edu/xw7/

Indiana University (Lead), Purdue, Carnegie Mellon, Duke, Yale, Penn State, Spelman, Ohio State, University of Illinois

NSF News

NSF announces awards to advance cybersecurity efforts

August 1, 2022

Cybersecurity is critical to safeguarding infrastructure, keeping supply chains moving, and ensuring privacy in cloud computing and health care. Adapting to ever changing threats requires cutting-edge research and transformative solutions.

The U.S. National Science Foundation is pleased to announce an investment of $25.4 million to advance ambitious research and center-scale projects in cybersecurity and privacy.

"The Secure and Trustworthy Cyberspace program is one of NSF's largest research programs, recognizing the criticality of cybersecurity and privacy to the nation's economy and to citizens," said NSF Director Sethuraman Panchanathan. "These investments support cybersecurity research across the country that can be translated into solutions that improve our quality of life."

IU cybersecurity researchers awarded multi-institutional NSF grants to protect data, user privacy

Indiana University will lead center focused on distributed confidential computing

FOR IMMEDIATE RELEASE | Aug. 4, 2022

Center for Distributed Confidential Computing

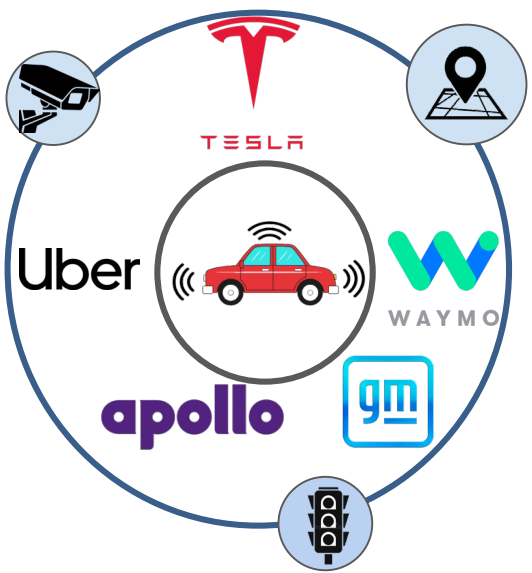
Led by Indiana University, this project will use the "trusted execution environment" hardware capability in modern chips to run secure computation in a way that can't be compromised by malicious software across distributed computing systems such as cloud computing environments. Researchers will work to provide solutions for data in use such as training machine learning models on private data, across cloud and edge systems. Indiana University will be joined by Purdue University, Penn State, Carnegie Mellon University, The Ohio State University, Spelman College, Duke University and Yale University will participate.

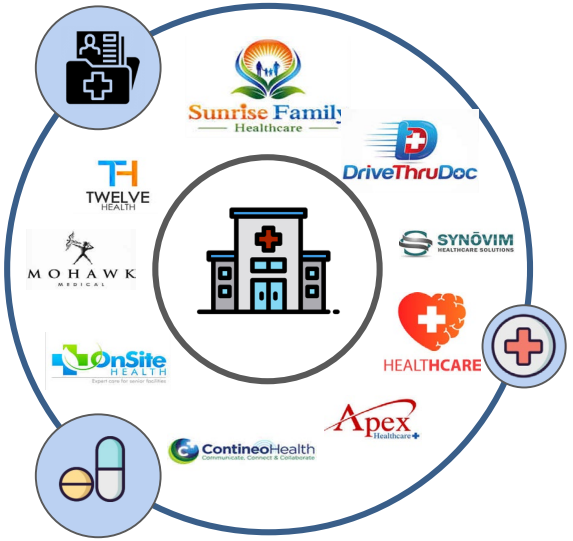
Learn more about the Secure and Trustworthy Cyberspace program and visit nsf.gov.

The National Science Foundation has made 9 million investment to build Center for Distributed Confidential Computing (CDCC), the first of its kind, to protect data-in-use in cloud-edge environments

Grand Challenge: How can we protect Data-in-Use at scale?

CDCC

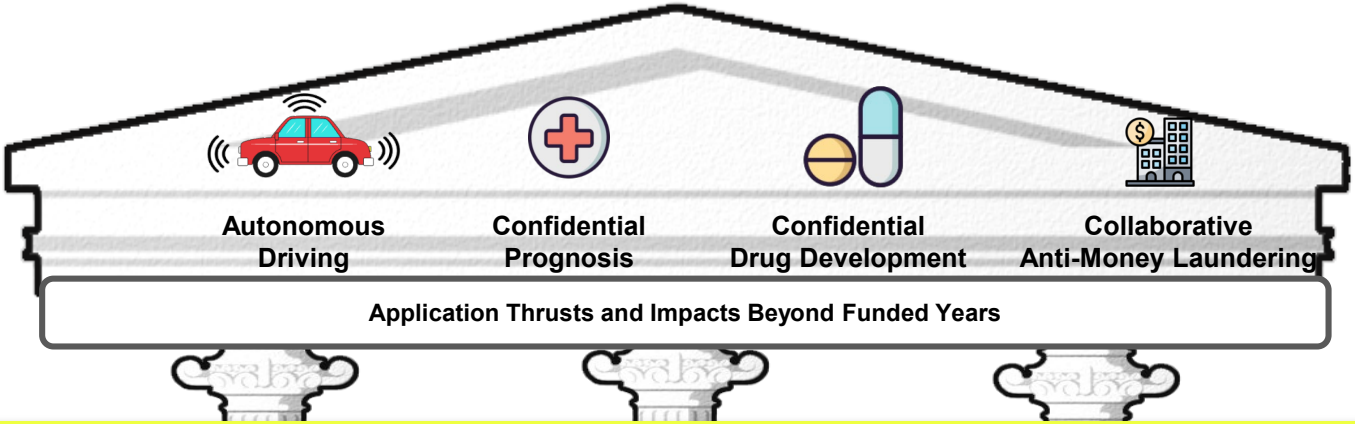




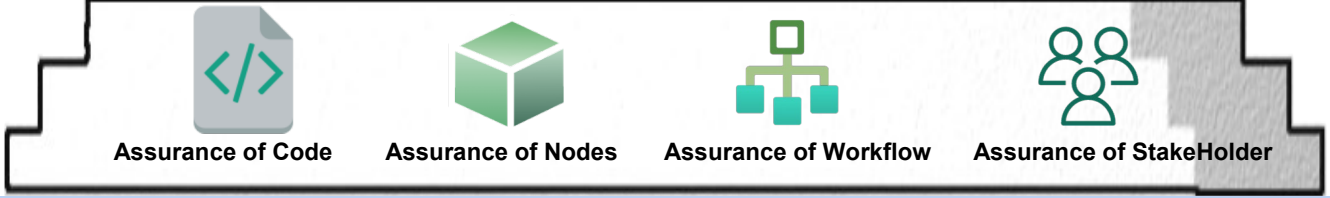
How can partners use data without disclosing it to unauthorized parties? How can this be done in a practical way?

Vision of Center for Distributed Confidential Computing

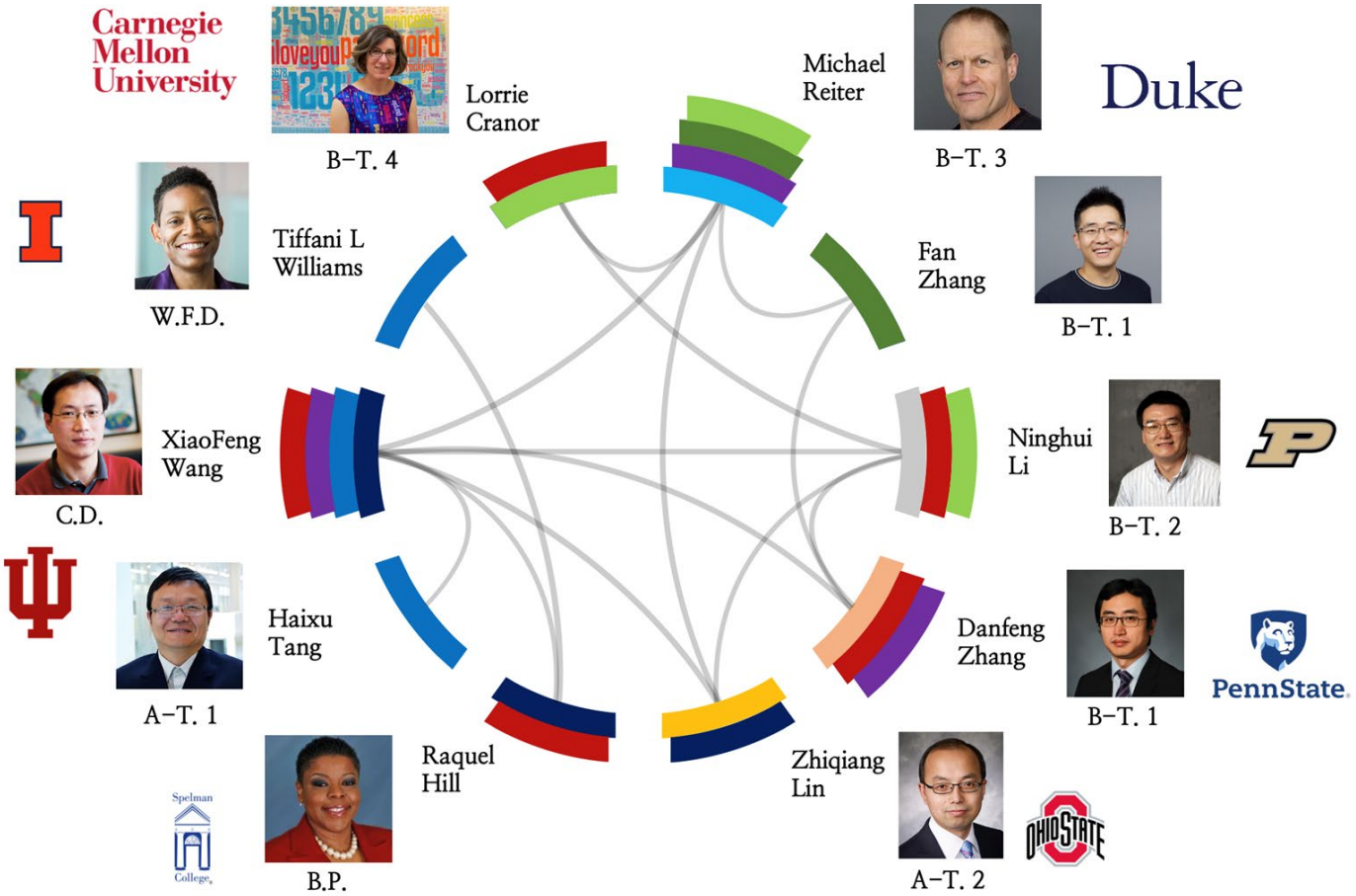
CDCC



Enable scalable, practical, verifiable and usable data-in-use protection to help maintain US leadership in AI and data science



Organization and Key Personnel: Academia



Fundamental Research Partners

Confidential Cloud

Ian Molloy
Department Head
IBM

Logan Ding
Applied Science Manager
AWS

Weidong Cui
Partner Research Manager
Microsoft Research

Jon McCune
Software Engineer
Google

TEE Middleware

Chris Ramming
Senior Director
VMware

Jethro Beekman
VP Technology
Fortanix

TEE Manufacture

Mona Vij
Principal Engineer
Intel

Brent Hollingsworth
Director
AMD EPYC

Advisory Board

Collaborator

Use-Inspired Research Partners

Disease Prognosis and Drug Development

Umberto Tachinardi
CIO
Regenstrief Institute

Barry Bunin
CEO & President
C.D.D.

Kun Huang
Investigator
Regenstrief Institute

Hua Xu
Working Group Chair
OHDSI consortium

IoT and Autonomous Driving

Gen Nishida
Senior Software Engineer
Cruise, GM

Ashish Kundu
Head of Cybersecurity R&D
Cisco

Hongxia Jin
Vice President
SamsungResearch

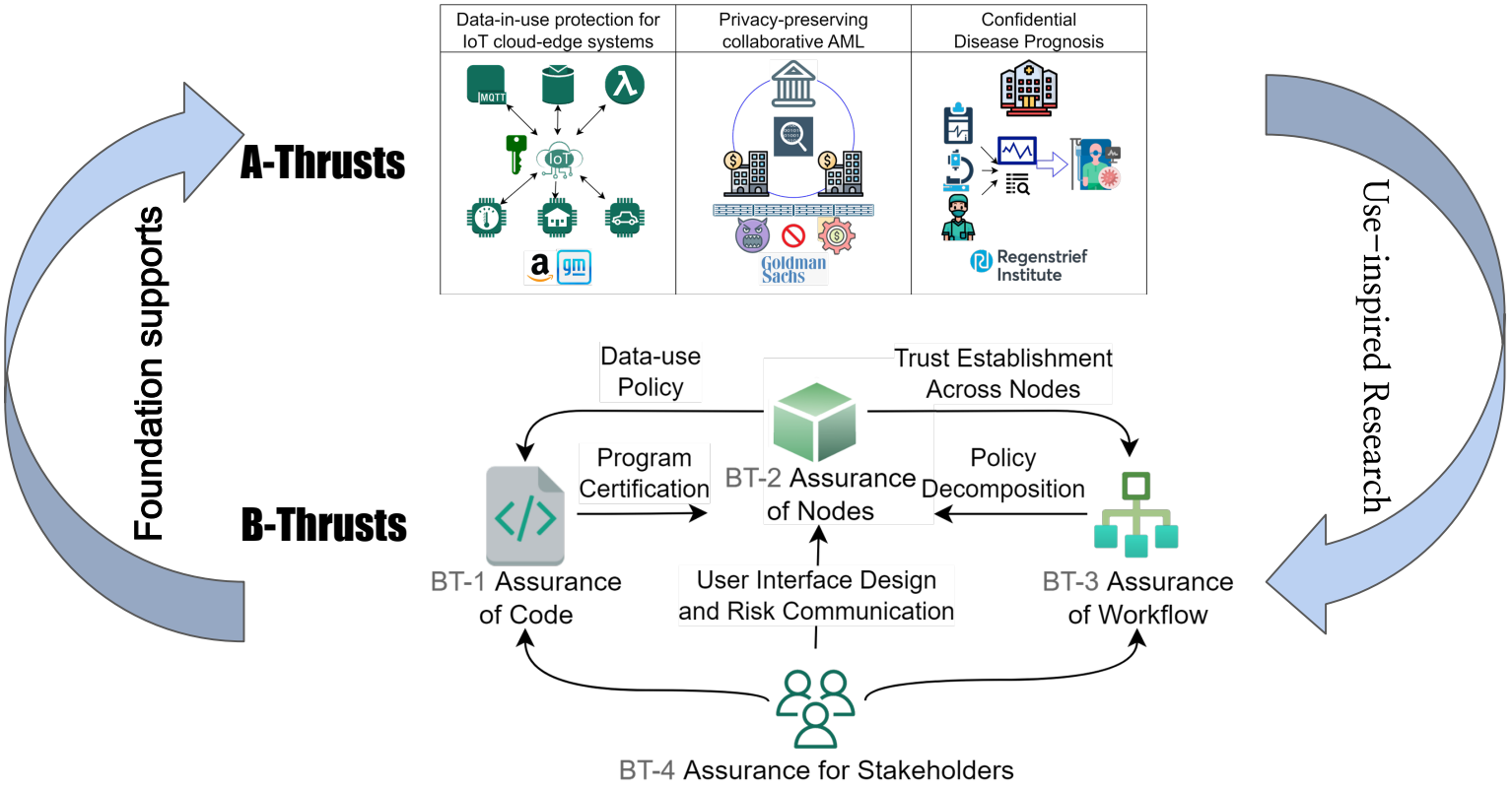
Anti - Money Laundering (AML)

Christos Boutsidis
Vice President
Goldman Sachs

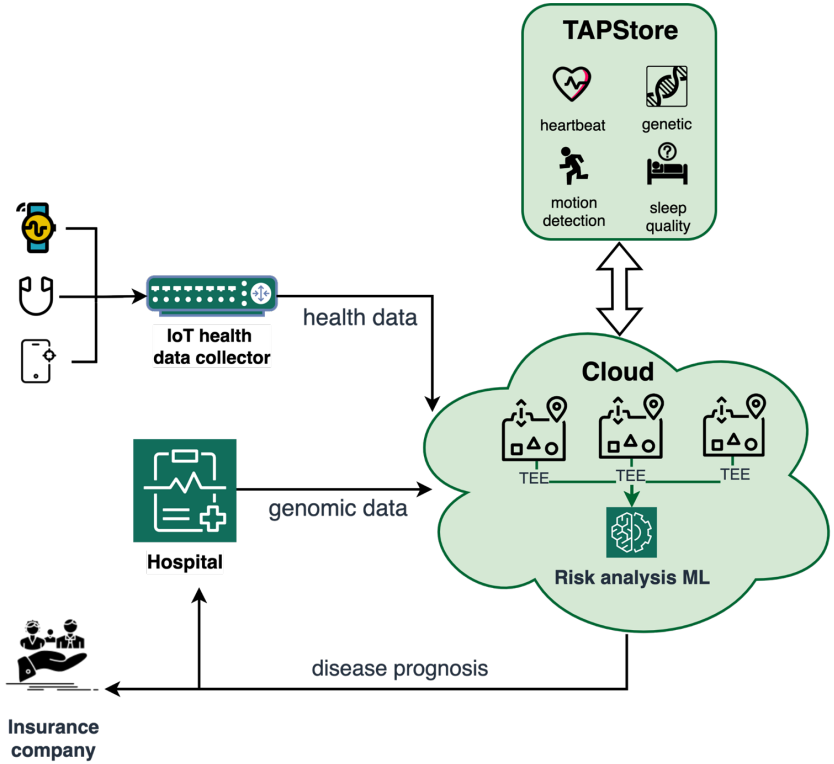
Advisory Board

Collaborator

Rationale and Synergies of Research Activities

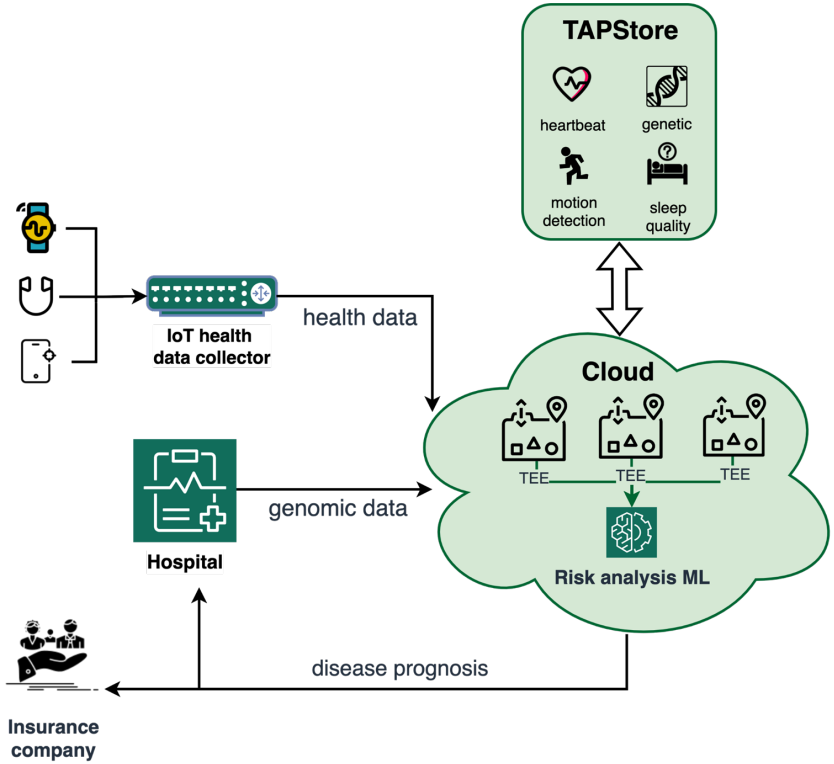


USE Case: Confidential Disease Prognosis



- **Goal: To predict the risk of critical clinical conditions from health and genomic data**
  - An application of distributed AI
  - An evolving process with new services continuous emerging

USE Case: Confidential Disease Prognosis



- **The patient's expectation for data protection**: All patient data and their derivatives will only go back to parties authorized by the patient (e.g., the hospital)

Use Case One: Confidential Disease Prognosis

- **Open problem 1:** how can enforceable policies on use of patient data be specified to meet the hospital's expectations?

Use Case One: Confidential Disease Prognosis

- **Open problem 1:** how can enforceable policies on use of patient data be specified to meet the hospital's expectations?
- **Open problem 2:** how can third-party prognosis programs be trusted to faithfully enforce the data-use policies?

Use Case One: Confidential Disease Prognosis

- **Open problem 1:** how can enforceable policies on use of patient data be specified to meet the hospital's expectations?
- **Open problem 2:** how can third-party prognosis programs be trusted to faithfully enforce the data-use policies?
- **Open problem 3**: how can the whole prognosis workflow offer adequate data protection at runtime?

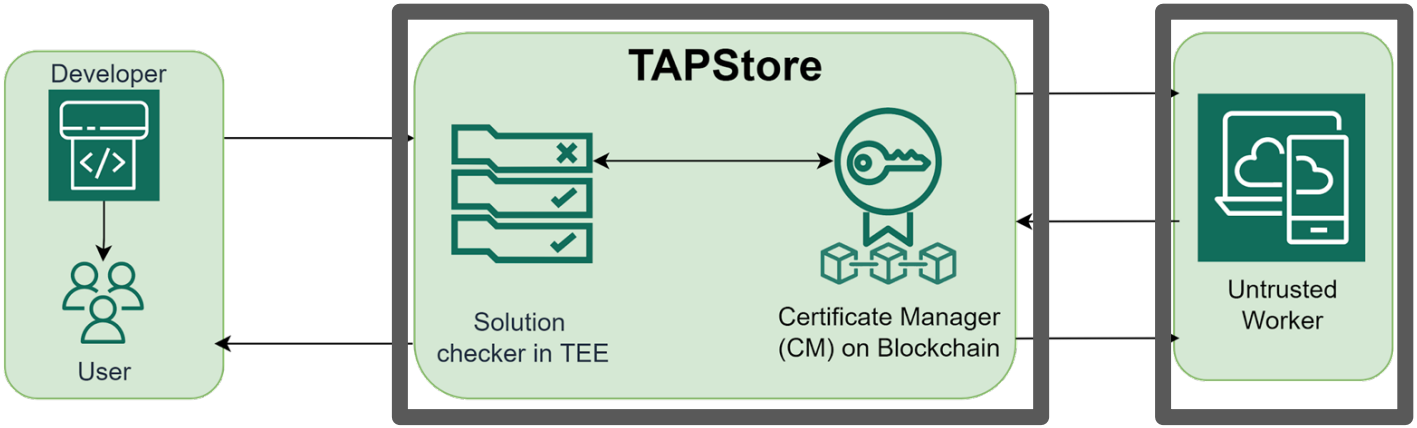
Use Case One: Confidential Disease Prognosis

- **Open problem 1:** how can enforceable policies on use of patient data be specified to meet the hospital's expectations?
- **Open problem 2:** how can third-party prognosis programs be trusted to faithfully enforce the data-use policies?
- **Open problem 3**: how can the whole prognosis workflow offer adequate data protection at runtime?
- **Open problem 4**: what user-friendly support is needed for hospital staff to effectively use data protection techniques?

B-Thrust 1: Assurance of TEE Code (PSU, Yale, Duke, IU and Purdue)

TAPStore Workflow

- Open
- Decentralized



Innovations in program verification and blockchain

No Trusted Third Party

B-Thrust 2: Assurance for TEE Nodes (Purdue, OSU, Spelman, IU and PSU)

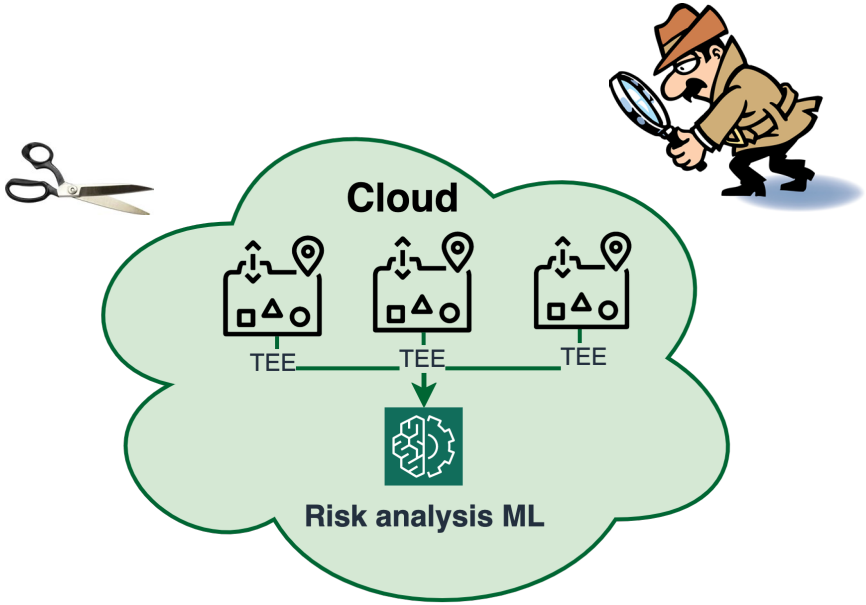
Strongly Enforced Dynamic Sticky (SEDS) Policies

- SEDS policy language design
- Enforcement at three levels: program, node (host), workflow
- Dynamically generated policies for data in motion and generated data
- SEDS policy analysis

Software on TEE Nodes

- Software for controlling I/O of programs that access data
- Scalable trust establishment between nodes
- Runtime and attestations across heterogeneous TEE nodes

B-Thrust 3: Assurance of TEE Workflow (Duke, IU, Spelman, PSU and Purdue)



SEDS Policy Decomposition

- Workflow policy breakdown to each TEE node

Leak Control on DCC Workflow

- Workflow visible leaks
- Topology leaks

Elastic DCC Support

B-Thrust 4: Assurance for Stakeholder (CMU, Duke and Purdue)

Stakeholders' Requirements and Preferences

- Understand requirements/preferences
- Guide the development of DCC technologies

Risk Communication with Stakeholders

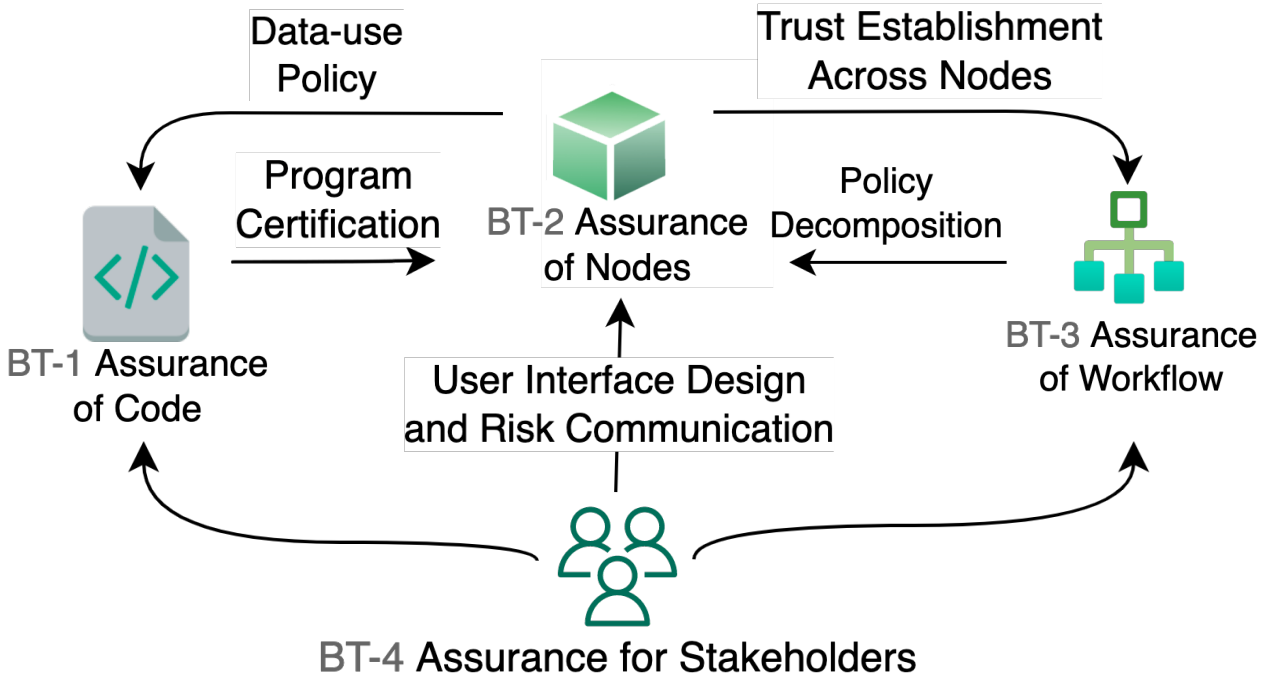
- Develop risk- communication techniques

Usability of the DCC Technologies

- Design, implement and evaluate UIs for DCC technologies

Rationale and Synergies of Research Activities

B-Thrusts



USE Case: Confidential Disease Prognosis

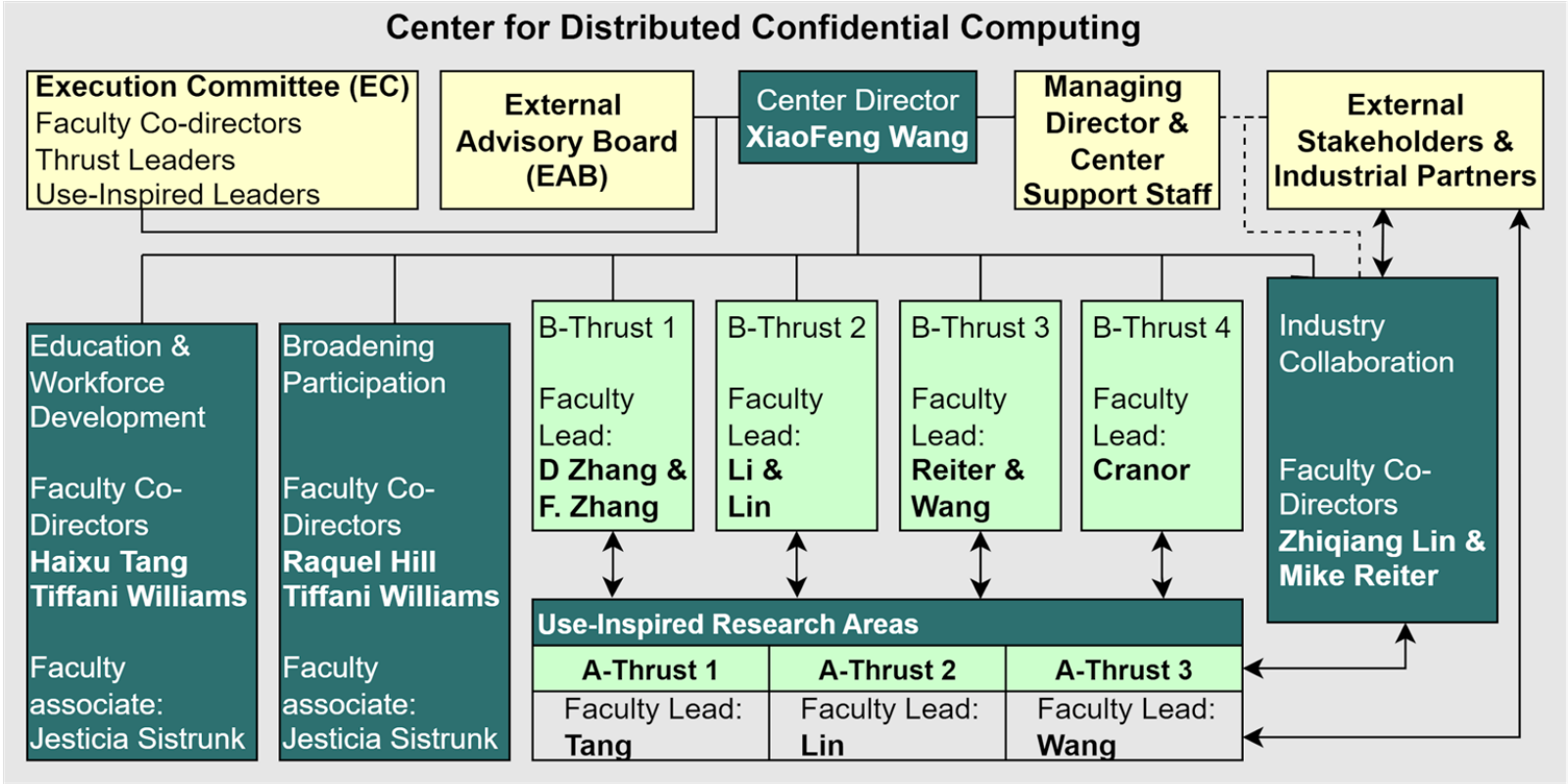
- **Open problem 1:** how can enforceable policies on use of patient data be specified to meet the hospital's expectations?
  Solution: A-Thrust 1 + B-Thrust 2
- **Open problem 2:** how can third-party prognosis programs be trusted to faithfully enforce the data-use policies?
  Solution: B-Thrust 1+2
- **Open problem 3**: how can the whole prognosis workflow offer adequate data protection at runtime?
  Solution: B-Thrust 2+3
- **Open problem 4:** What user-friendly support is needed for hospital staff to effectively use data protection techniques?
  Solution: B-Thrust 1+2+3+4

Rationale and Synergies of Education and Outreach

- **CDCC will commit to cybersecurity workforce development and diversification:**
  - Developing the first confidential computing curriculum and make it a **model** worldwide
  - Expanding access of under- represented groups to DCC through an ambitious BP plan
  - Collaborating with industry to transfer knowledge and develop a qualified workforce

- **Our research, education and outreach activities will induce a domino effect:**
  - More diverse and better trained workforce leads to better R&D
  - Leads to more technology transfer and better products
  - Causes more excitement in DCC and attracts more involvement in related R&D activities

CDCC Management and Integration



Expected Achievements

- DCC foundation to enable practical data - in- use protection for big data analytics and AI computing
- Practical advances in use cases validated by collaborators
- Technology Transfer and Standardization
  - Open- Source Projects and Working Group on healthcare data- in- use protection
- Impacts on academia: influential papers
- Adoption of DCC courses across partner universities and beyond
- Model to create broad and diverse DCC communities

Additional Supports for Impact Amplification

- Hardware- specific customization
  - Utilization of unique hardware features (e.g., Intel's IPU) for scalable DCC
  - Analysis on unique privacy risks of specific hardware platforms
  - Customization of techniques to various platforms
- Incorporation of software - based DCC
  - Seeking cost effective DCC solutions involving both TEE and crypto solutions (FHE, SS, MPC etc.)
- Supports for other application domains
  - E.g., Trustworthy Data Center

Join Us as a Partner (Contact: xw7@iu.edu)

- Research and education agendas
  - In- center personnel to guide related research and other activities
  - Invitation to the annual industry meeting
  - Seat in the EAB
- Joint research and technology transfer
  - Sponsoring separate projects with Center researchers
  - Customizing technologies to be developed on the partner's hardware platform
  - Joint basic research, publications and proposals
  - Early access to unpublished research
- Workforce development
  - Targeted curricular development providing future employees with focused skills and minimal learning curves
- Recruitment
  - Priority in accessing well- trained students/postdocs for interns and recruitment
  - Assistance in finding suitable faculty experts for consultancy
  - Support in diversifying workforce through the center's out- research to under represented population
- Publicity
  - Featured place on NSF CDCC website and publications
  - Invitation to deliver technical presentations in the CDCC seminar series
  - Opportunities to sponsor customized workshops at member sites

Center for Distributed Confidential Computing (CDCC)

Thank You