

What We Can't See, We Can't Fix

**The Quest for Transparent Broadband Data
Infrastructure**

Arpit Gupta

Associate Professor, UC Santa Barbara

Faculty Scientist, Berkeley Lab

Fellow, Benton Institute

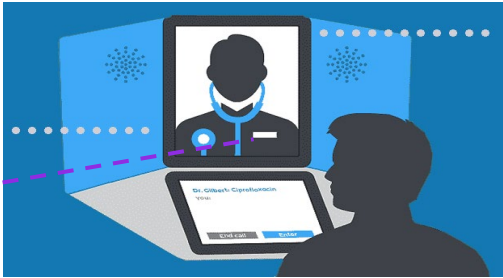
The Fundamental Paradox

Education

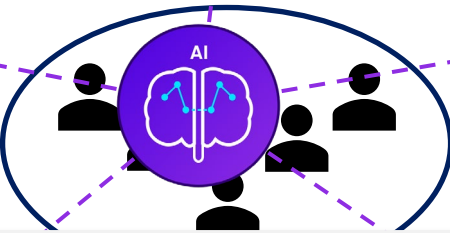
E-Gov



Healthcare



Work



Finance

This transformative Internet does not work for ALL!



Goal: The Pursuit of “Universal Access”

\$100B+ spent in the US in the past 30 years

Program	Funding	Description
Broadband Equity, Access and Deployment (BEAD) Program	\$44 Billion	Fund new infrastructure deployments to ensure high-speed internet for all
Connect America Fund (CAF)	\$10 Billion	Subsidize new infrastructure in hard-to-serve areas

**Policymakers need data to assess true state of broadband
Availability, Affordability, Adoption, and Quality**

Problem: Lack of Transparent Data Infrastructure

Internet Service Providers

Policymakers

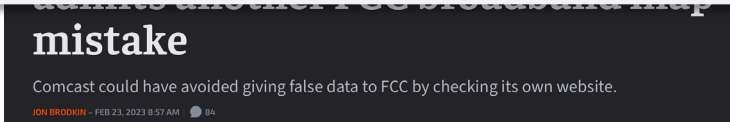


Unreliable &
Noisy

Internet Availability, Quality, and Cost



How to reduce dependence on **unreliable and noisy** (self-reported) data from ISPs?



Observation: Key Information is Publicly Available

Accurate information readily available to potential customers at street address granularity



Availability [5663 Cielo Ave, #B, 93117](#) [Edit address](#)

Unlock extra savings on select Cox Internet plans and experience unbeatable 5G reliability on-the-go when you add Cox Mobile after purchase. [Find out how](#)

GET HIGH-SPEED INTERNET AT AN AFFORDABLE PRICE FROM A LEADING INTERNET PROVIDER

Delve into a realm of fast and consistent connectivity without breaking the bank. Our competitive pricing coupled with unparalleled customer service makes Cox the go-to internet provider for budget-friendly high-speed internet solutions.

SAVE \$15.00/mo. with Cox Mobile INCLUDES Panoramic Wifi equipment

Quality

100 Mbps
Max download speed

Go Fast
Great for browsing the web, social media, email, and more

Cost

\$35.00/mo
with Cox Mobile, for 24 mos. No annual contract or cancellation fees.

\$50/mo
without Cox Mobile, no annual contract or cancellation fees.

[Order now](#)
[Customize plan](#)
[Plan details](#)

The screenshot shows a website interface for Cox Internet. At the top, the word 'Availability' is followed by a house icon and the address '5663 Cielo Ave, #B, 93117', which is circled in red. To the right is an 'Edit address' link. Below this is a green banner with a mobile phone icon and text about savings and 5G reliability, with a 'Find out how' link. The main content area has a heading 'GET HIGH-SPEED INTERNET AT AN AFFORDABLE PRICE FROM A LEADING INTERNET PROVIDER' and a paragraph of descriptive text. Below this is a section with two columns. The left column is titled 'Quality' and features '100 Mbps' with a speed icon and 'Max download speed', circled in red. The right column is titled 'Go Fast' and describes the service's use cases. Below these is a 'Cost' section with two pricing options: '\$35.00/mo' (with Cox Mobile) and '\$50/mo' (without Cox Mobile), both circled in red. On the far right are three buttons: 'Order now', 'Customize plan', and 'Plan details'.

Solution: Broadband-Plan Querying Tool (BQT)

Automates querying ISPs' web interfaces to extract advertised broadband speeds and prices at street addresses granularity



Enabled querying 1M+ addresses for 100+ ISPs

Impact: BQT filled a Critical Data Gap

SIGCOMM'24 RTF Applied Networking Research Prize'25

The Efficacy of the Connect America Fund in Addressing US Internet Access Inequities

Haarika Manda*, Varshika Srinivasavaradhan*, Laasya Koduru, Kevin Zhang, Xuanhe Zhou, Udit Paul†, Elizabeth Belding, Arpit Gupta, Tejas N. Narechania‡
UC Santa Barbara † Ookla Inc. ‡ UC Berkeley

SIGCOMM'23 SIGCOMM Doctoral Dissertation Award 2024

Decoding the Divide: Analyzing Disparities in Broadband Plans Offered by Major US ISPs

Udit Paul*, Vinothini Gunasekaran*, Jiamo Liu*, Tejas N. Narechania§,
Arpit Gupta*, Elizabeth Belding*

Amicus Curiae Brief

No. 23-1127

In the
Supreme Court of the United States

WISCONSIN BELL, INC.,

Petitioner,

v.

UNITED STATES OF AMERICA EX REL. TODD HEATH,

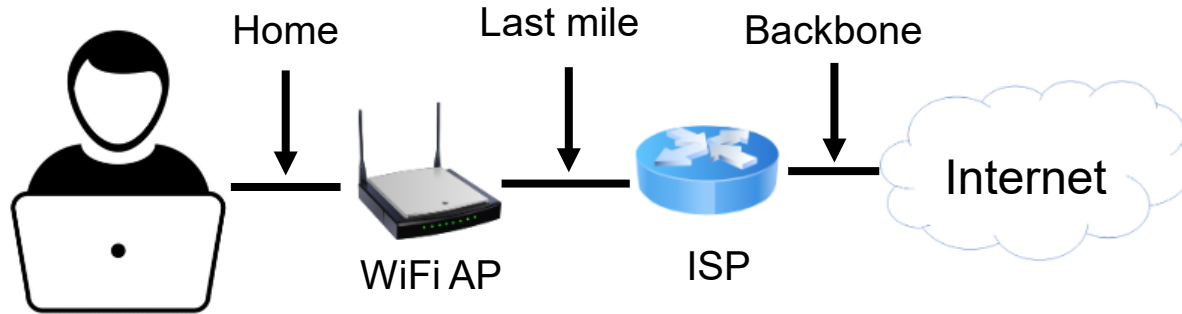
Respondent.

On Writ of Certiorari to the

The Next Frontier

Reduce the technical debt of extending & operating BQT

Missing Piece: What about Broadband Quality Data?

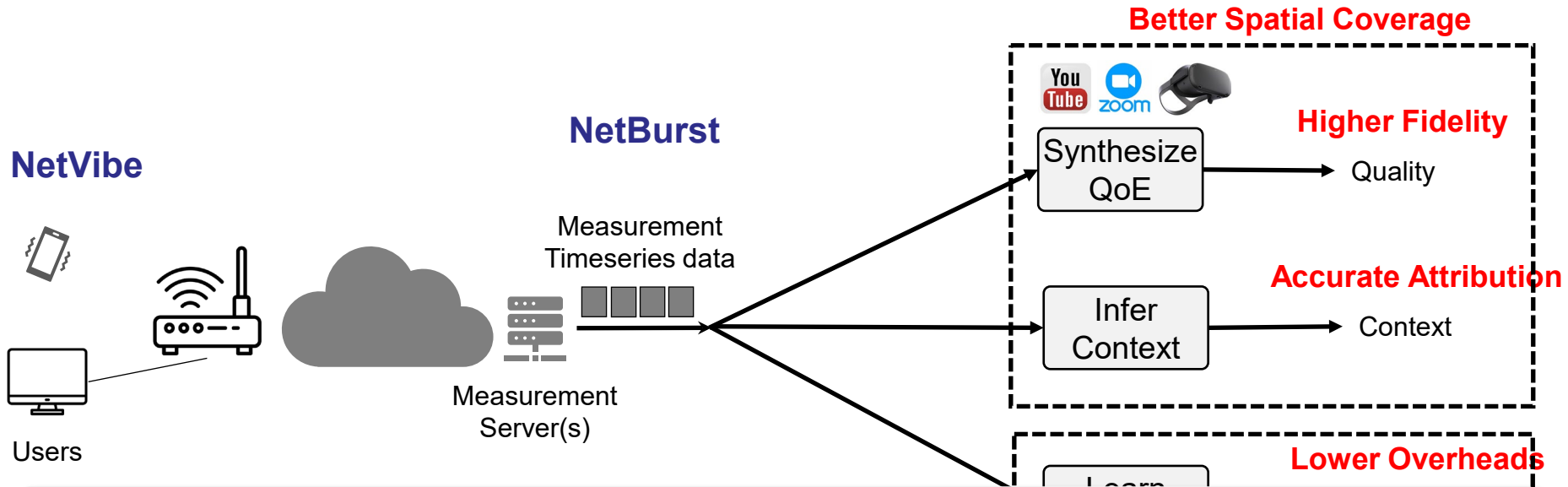


For every residence (i.e., **dense spatial coverage**), periodically (i.e., **dense temporal coverage**)

- Report metric(s) that captures users' quality of experience* (**high-fidelity assessment**)

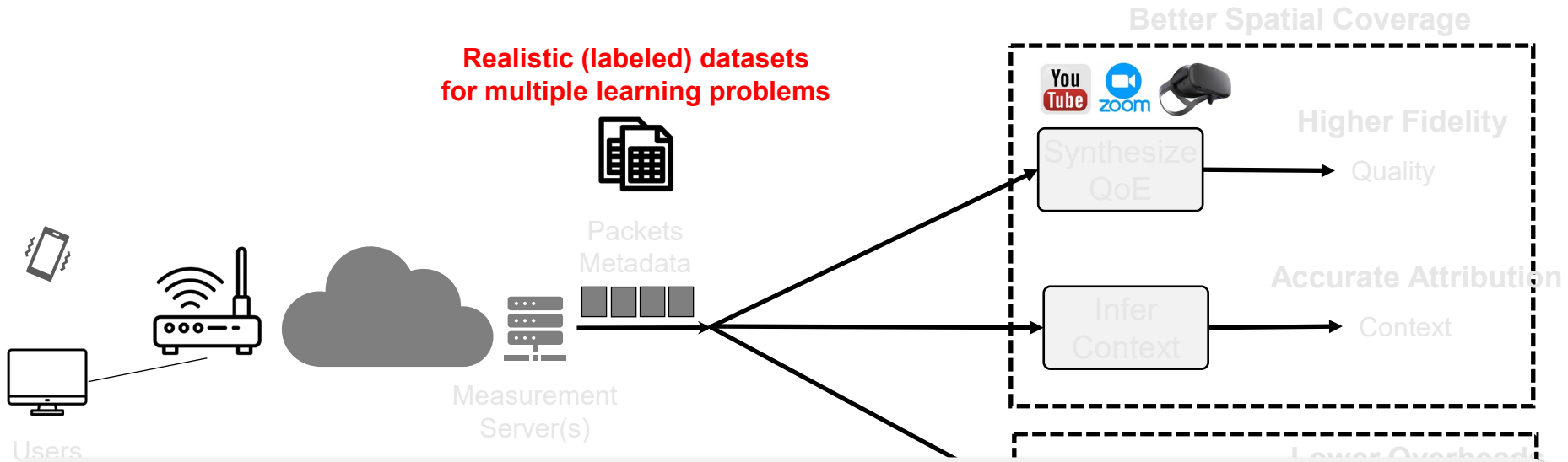
**Existing crowdsourced tools fail to bridge gap between
measured and experienced broadband quality**

Rethinking Broadband Quality Measurements



Learn data representations that enable **context inference & counterfactual models with low-cost measurements**

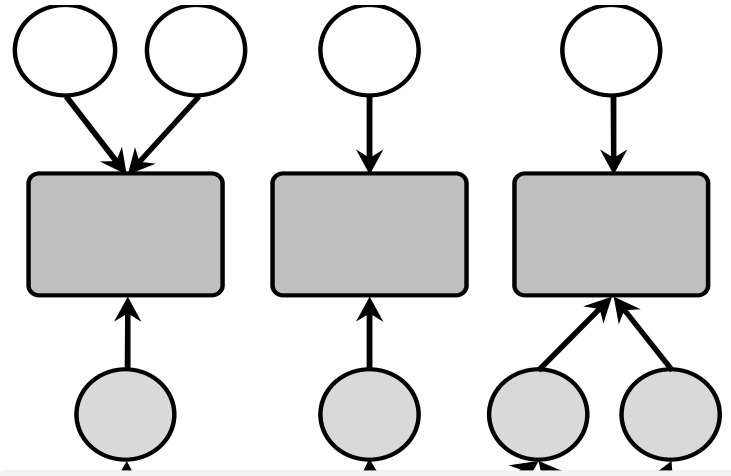
How to Enable ML-Powered Quality Measurements?



Programmatically generate realistic data realizing disparate intents with high fidelity

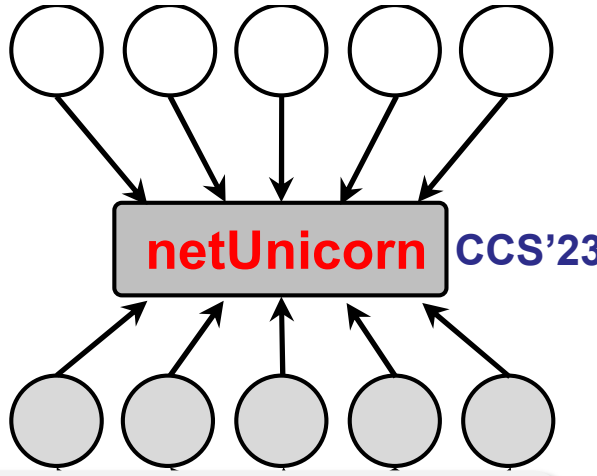
How to Fix an Inherently Fragmented Ecosystem?

Fragmented



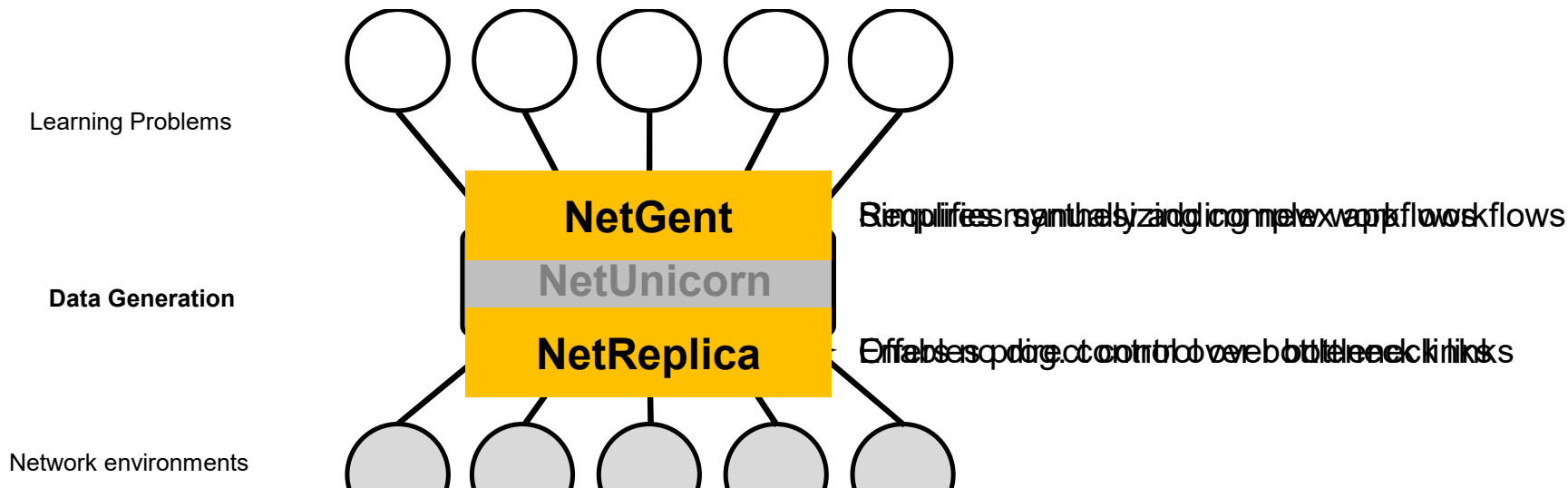
Thin Waist

Learning Problems
Data Generation
Network environments



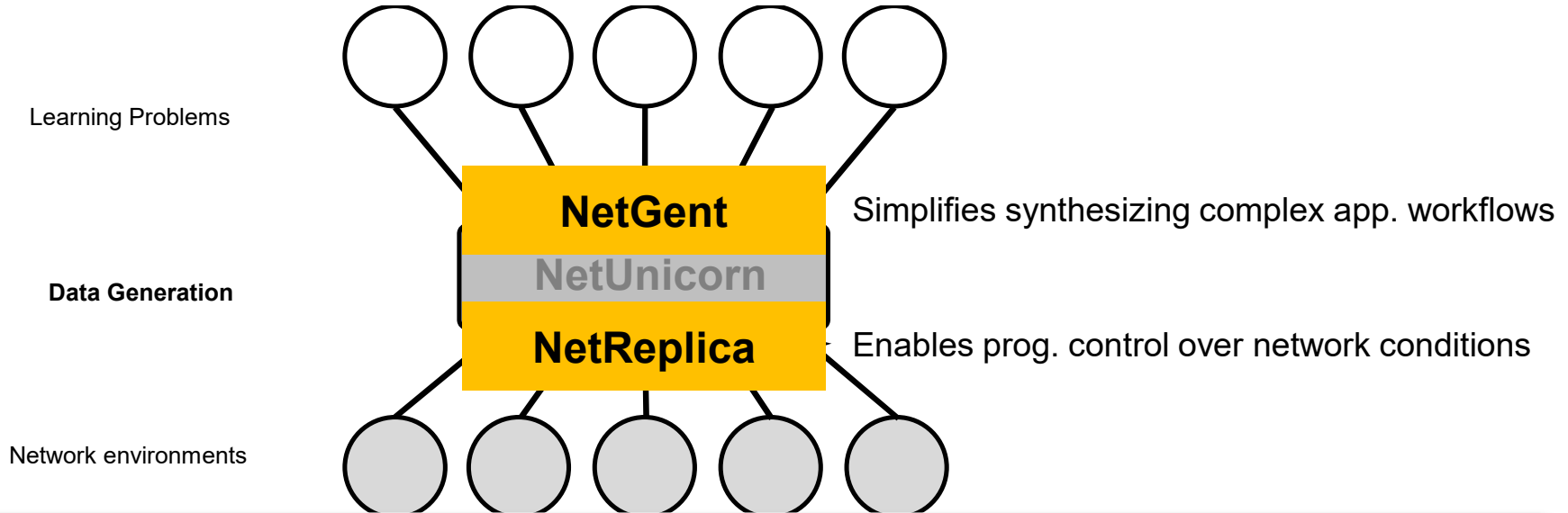
Disaggregates
intent from execution and execution from infrastructure

Shoring Up this Data Generation “Thin Waist”



Progressively **disaggregates** the data-generation substrate

A Programmable Data Generation Substrate



Enables generating traffic for 100s of applications across millions of disparate network conditions

Summary

This effort aims to develop **transparent broadband data infrastructure**—bridging the gap between required and available data in the public domain

- **Broadband Availability, Affordability, and Adoption**

Developed **BQT**, which enabled a unique infrastructure for broadband availability, affordability, and adoption data

Open problem: Scaling up the infrastructure with minimal cost and technical debt

- **Broadband Quality**

Embarked on a path to bridge the gap between experienced and measured broadband quality

Requires: Fundamental innovations in high-fidelity data generation, representation learning, and machine learning