



**Boston University** Wheelock College of Education  
& Human Development



Software & Application  
Innovation Lab

# Open-Source Infrastructure for Student Learning Disability Identification and Treatment

Red Hat Research Interest Group  
meeting

# Our Team

PI: Hank Fien, PhD

Professor, BU Wheelock

Co-PI: Ola Ozernov-Palchik, PhD

Senior Research Scientist, BU Wheelock

Co-PI: Eshed Ohn-Bar, PhD

Assistant Professor, BU Electrical & Computer Engineering

Will Tomlinson, PhD

Director, Software & Application Innovation Lab (SAIL)

Arezoo Sadeghi

Software Engineer, Software & Application Innovation Lab (SAIL)

Greg Frasco

Associate Director of Engineering, SAIL

Catherine Chan-Tse

Senior Software Engineer

Sam Lindberg

Technical Support Engineer, Red Hat

# The Problem

literacy  
is a  
civil right

- Only 32% of 4th graders are proficient readers (NAEP, 2022)
  - **11%** of Students with disabilities
  - **10%** of ELL students
  - **17%** of Black students
  - **19%** of low income students
- 130 million adults with low literacy skills (U.S. Department of Education)
  - 54% of adults read below sixth-grade level

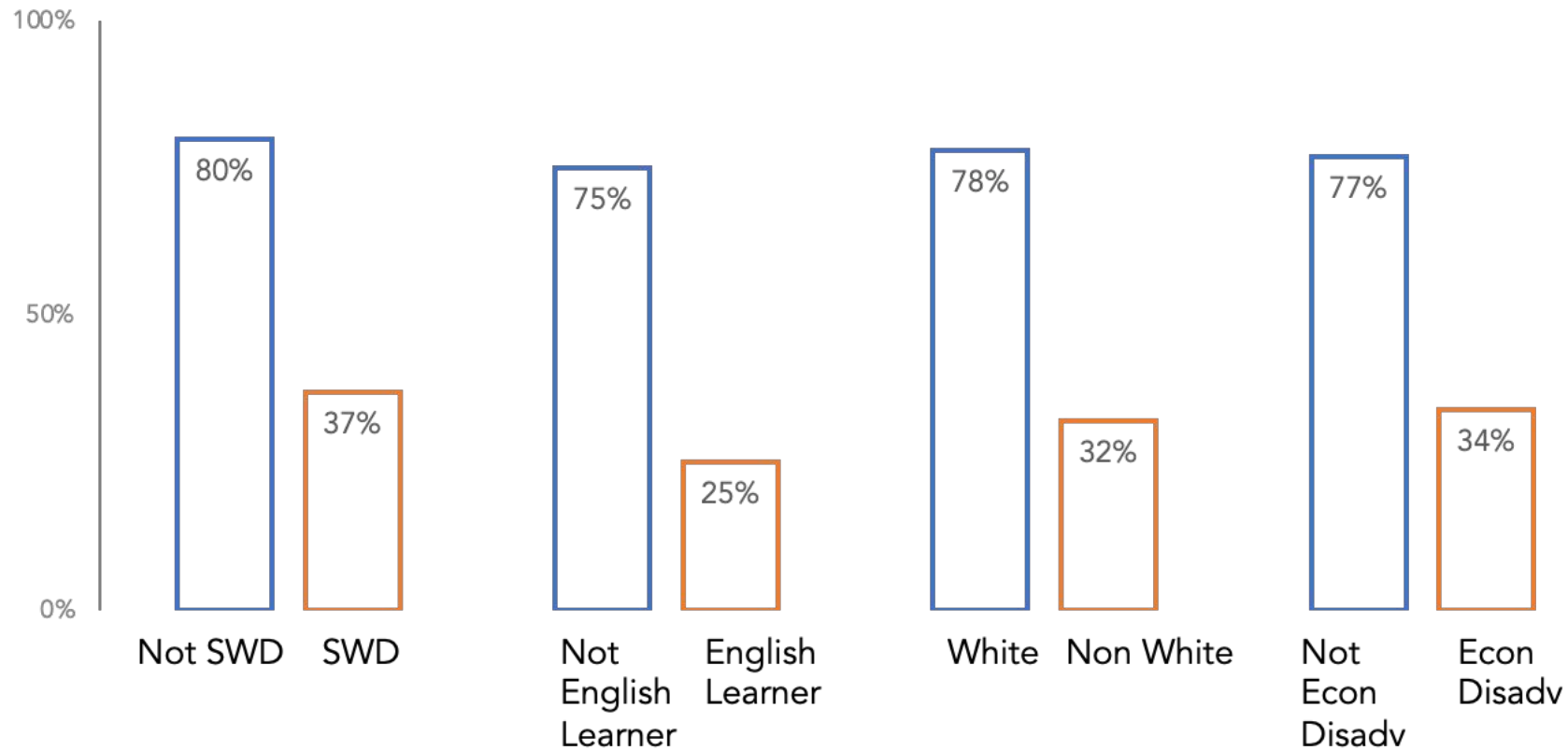
Disparities in wealth, employment, and virtually all other indicators of health and well-being (Kendi, 2020)

# The Problem

- 34% of 4th-grade children read below Basic, and 89% of students with disabilities scored below basic levels (NCES, 2022).
- There are historic dips in reading achievement for all school-aged children i (NAEP, 2022).
- Reading achievement gaps are significant for children with reading disabilities, low-income, and racial and ethnic minority students (NAEP, 2019).
- COVID-19 pandemic has worsened these reading gaps for marginalized and vulnerable student populations.

# Down the Road from BU...

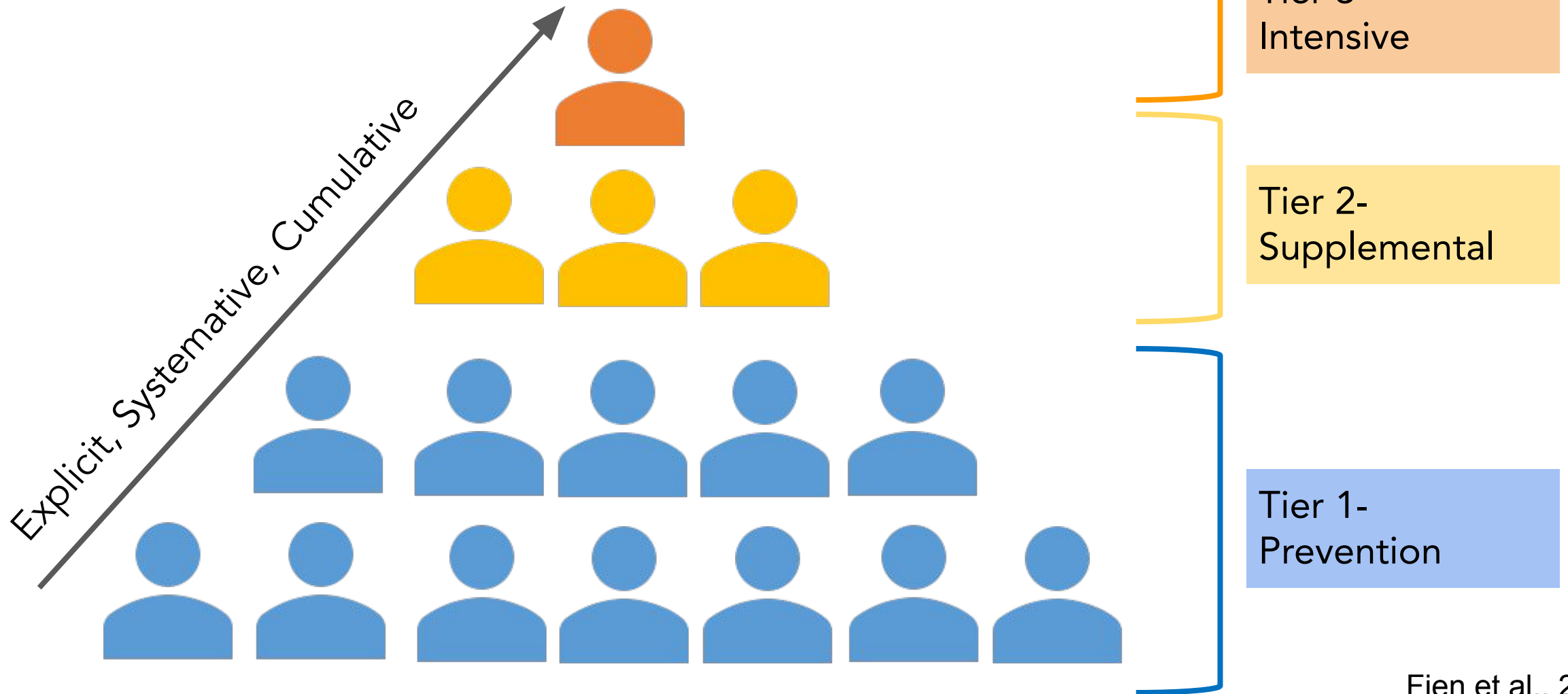
PSB Percent of Students Meeting & Exceeding Expectations (Grades 3-8 ELA)



SWD= students with disabilities

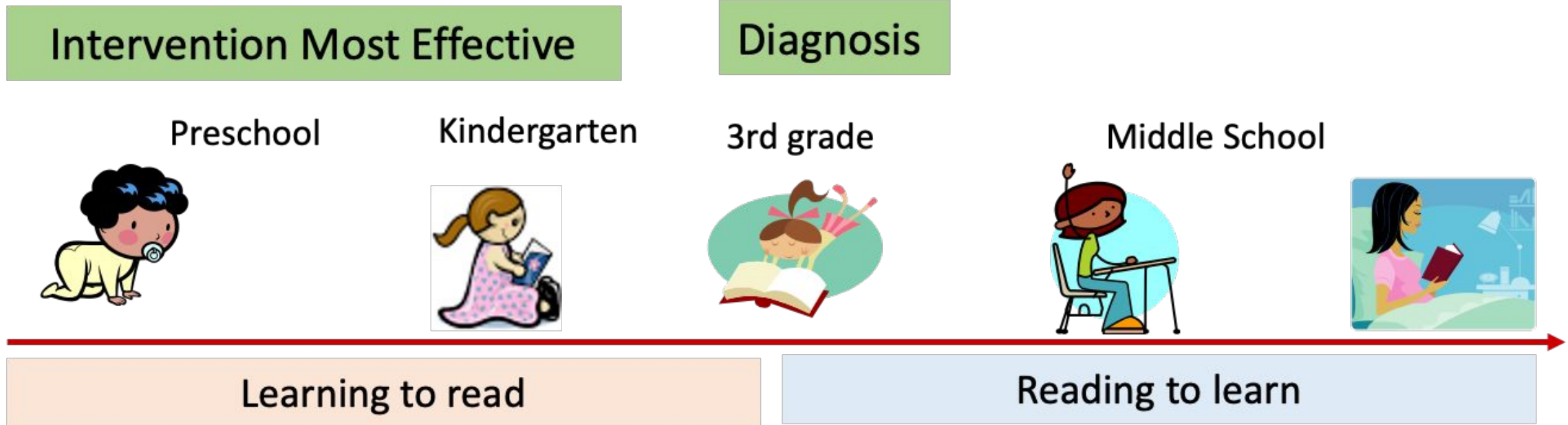
Econ Disadv= economically disadvantaged

# Evidence-based Solution



# Prevention Paradox

To date, the earliest that reading difficulties can be reliably diagnosed is in third grade, but interventions are most effective in KG-2nd grade



# Negative Implications



- 74% of children who are reading disabled in 3rd grade remain disabled by the end of high school
- Teachers, parents, and peers often misinterpret child's struggle to learn as negative attitude or poor behavior and decreased self esteem is often a result (Riddick et al., 1999)
- These negative experiences leave children with disabilities vulnerable to feelings of shame, failure, inadequacy, helplessness, depression, and loneliness (e.g., Valas et al., 1999)
- Less likely to complete high school or join programs of higher education and more likely to enter the juvenile justice system (Wagner et al., 1993)



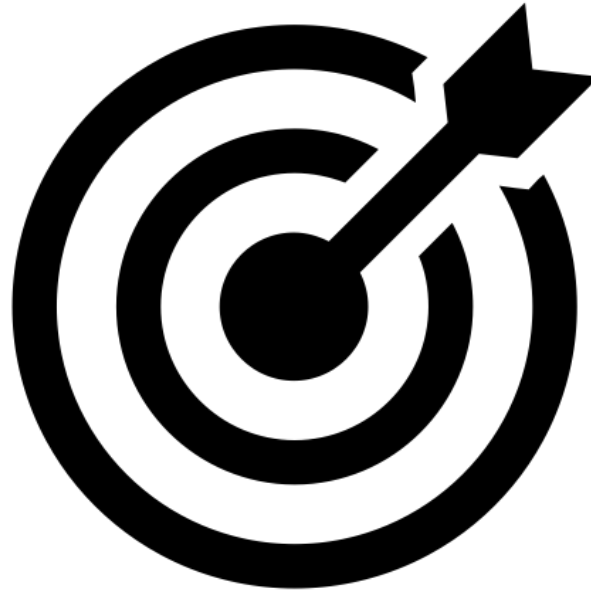
# The Need for Precision



Risk identification

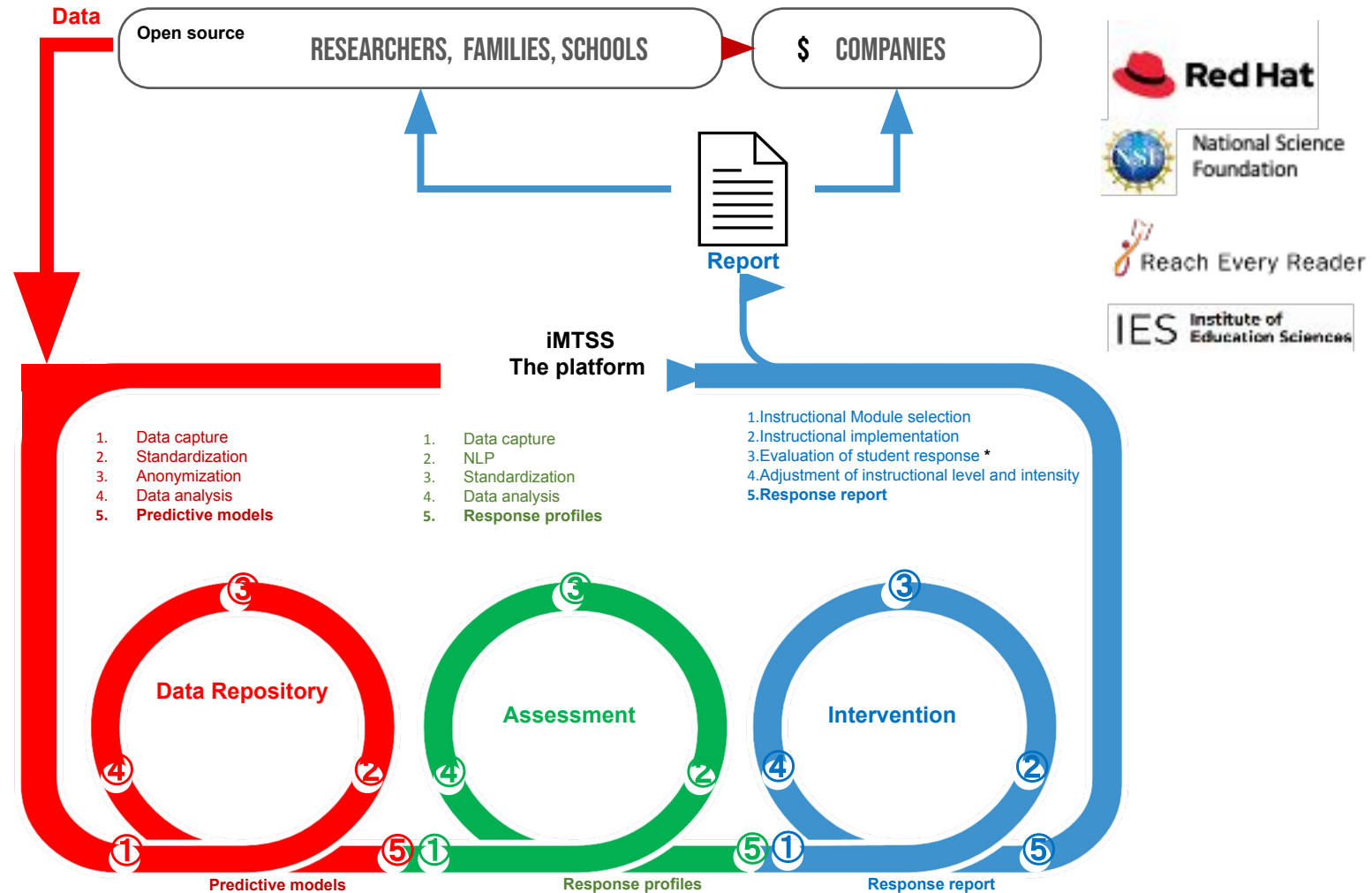


Response prediction



Big Data +  
Intelligent Modeling

# Digital Multi-Tiered System of Support (MTSS) Platform

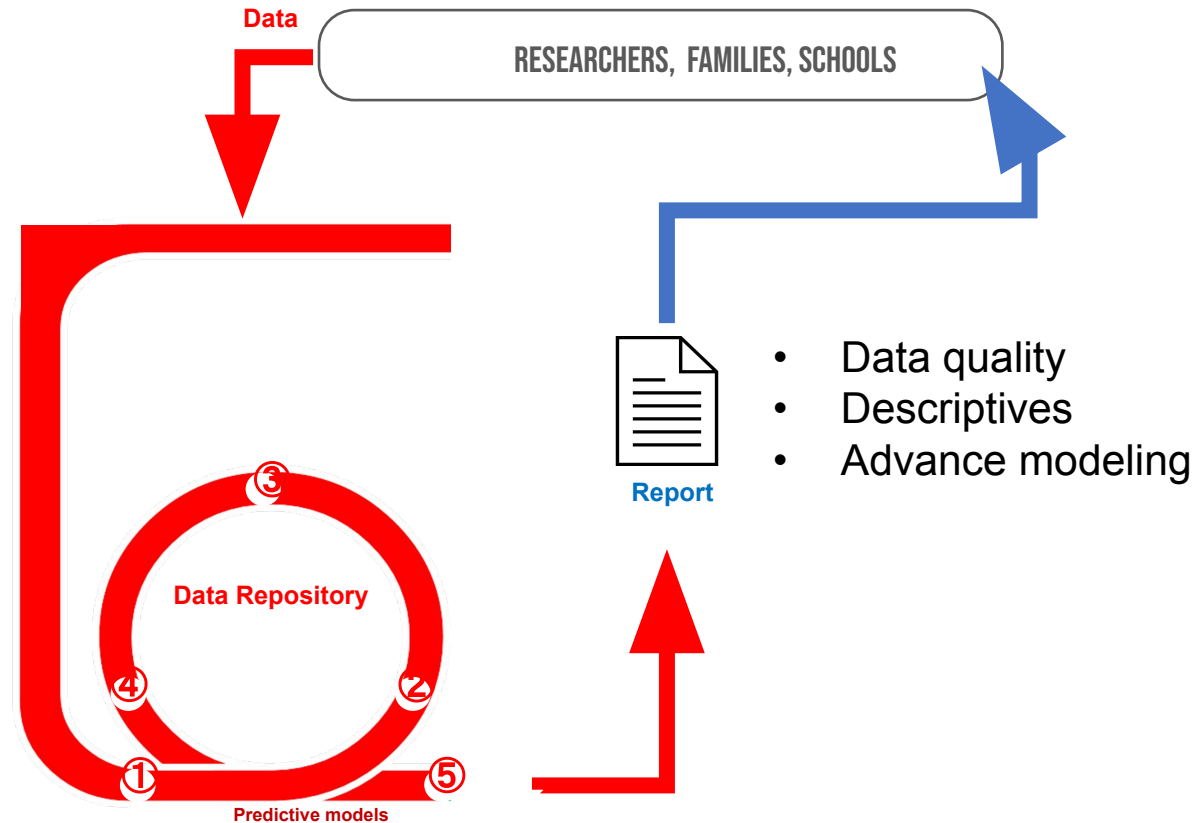


# “Intelligent” Educational Data Repository

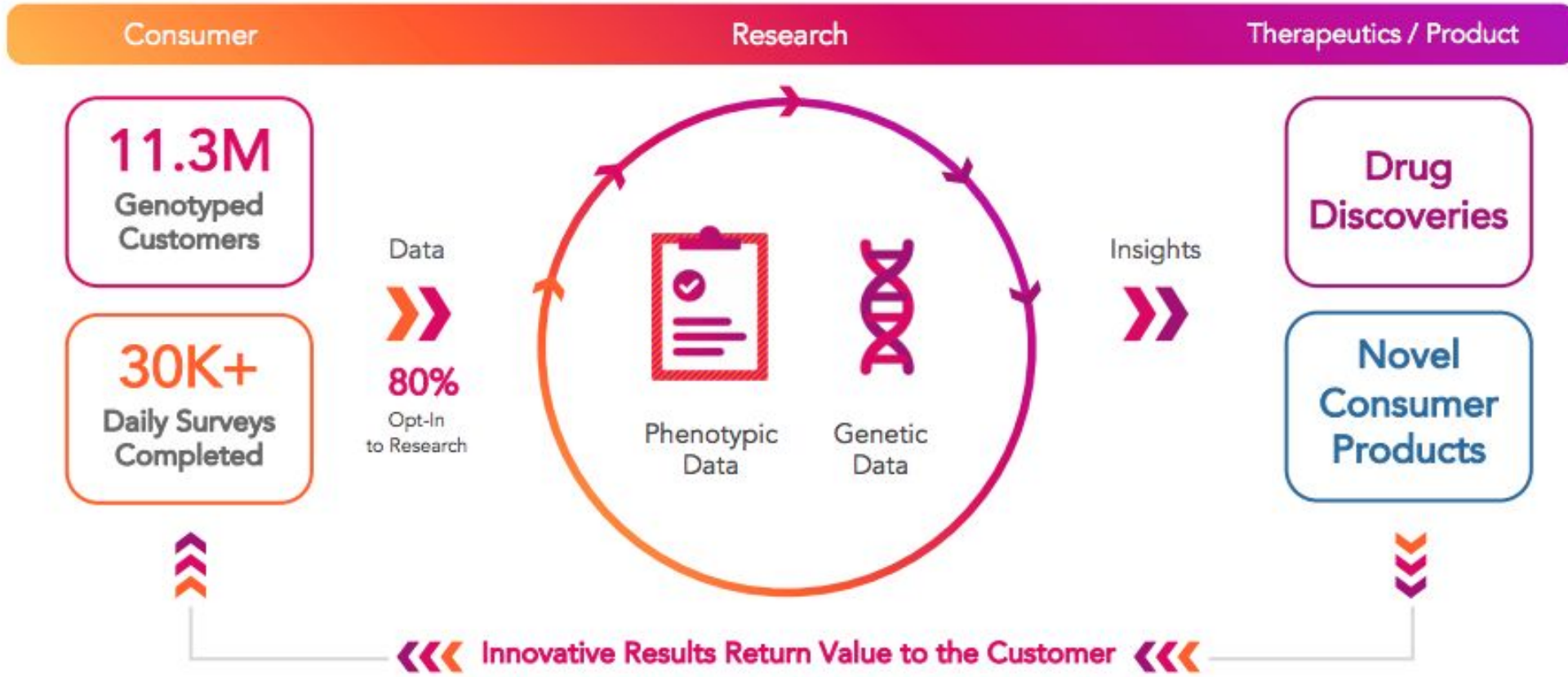


✓ Open source

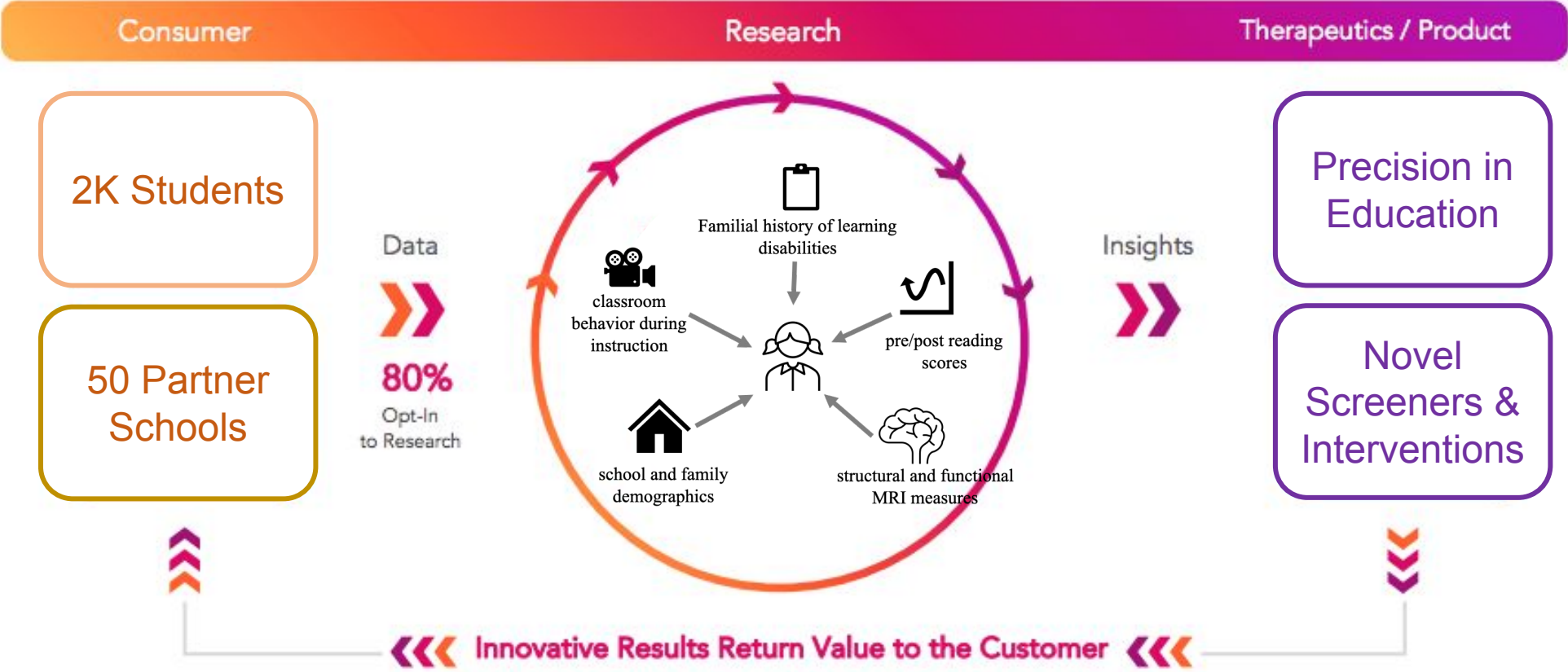
1. Data capture
2. Standardization
3. Anonymization
4. Data analysis
5. Predictive models



# 23andMe: Case study in Scalability



# Applying the Model to Educational Data



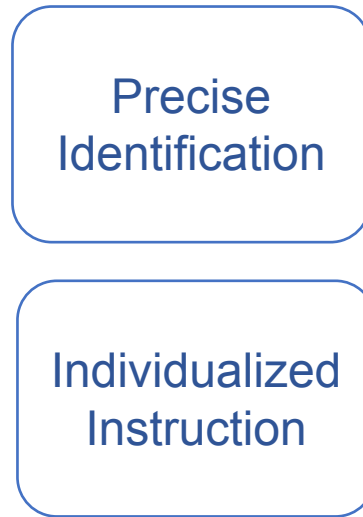
# Use Case: A Classroom Teacher



Teacher input data  
for individual student



Teacher receives a  
comprehensive risk  
profile of the student  
and  
recommendations  
for intervention



Student receives  
individualized  
interventions in  
appropriate dosages



Student's  
educational  
outcomes are  
optimized

# Challenges

Challenge 1. Data privacy



# Challenges

Challenge 1. Data privacy

Challenge 2. Flexibility across data types/users



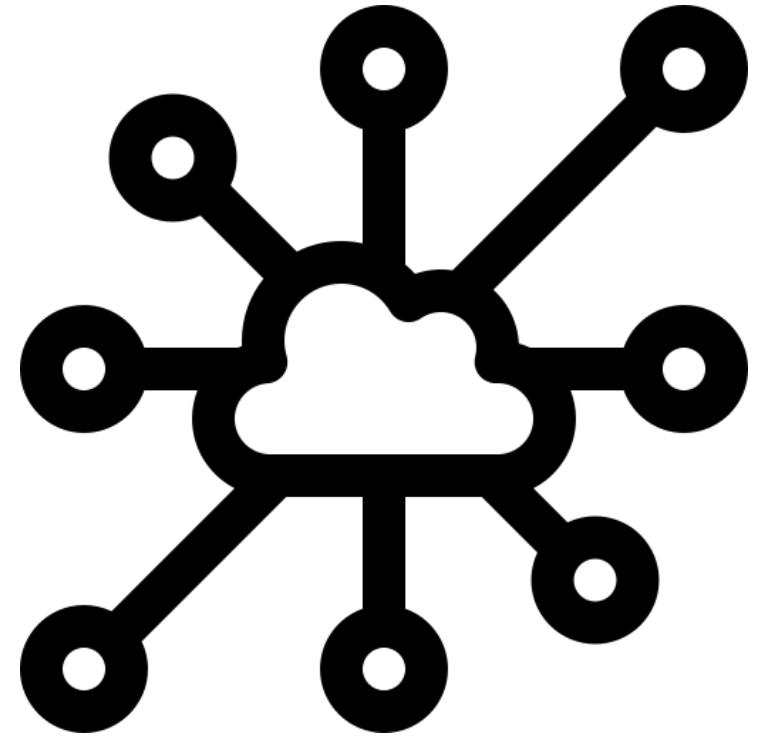
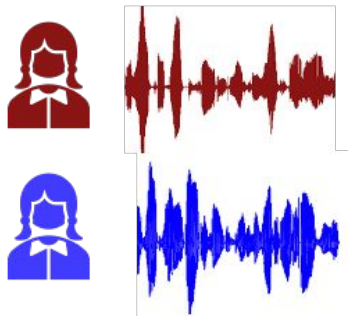


# Challenges

Challenge 1. Data privacy

Challenge 2. Flexibility across data types/users

Challenge 3. Scalability to large datasets



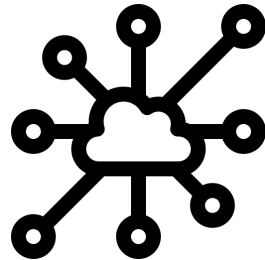
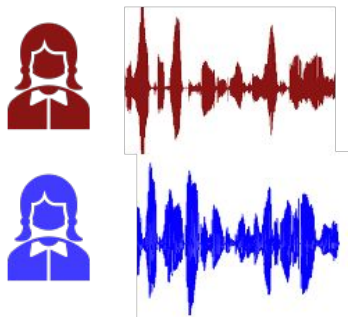
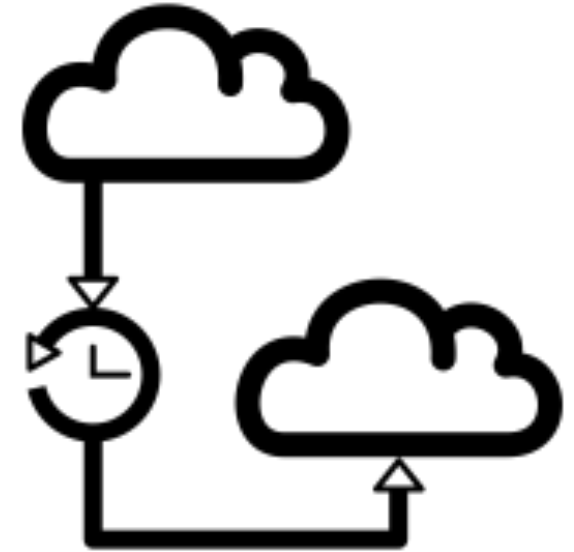
# Challenges

Challenge 1. Data privacy

Challenge 2. Flexibility across data types/users

Challenge 3. Scalability to large datasets

Challenge 4. Latency (in real-time)



# Challenges

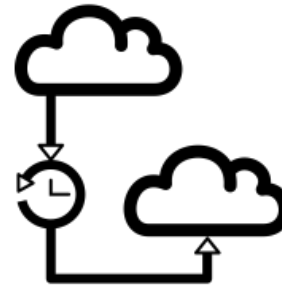
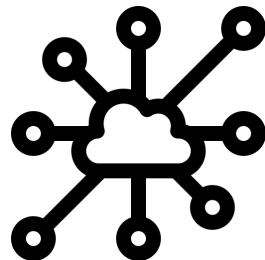
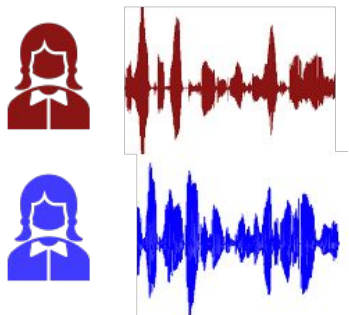
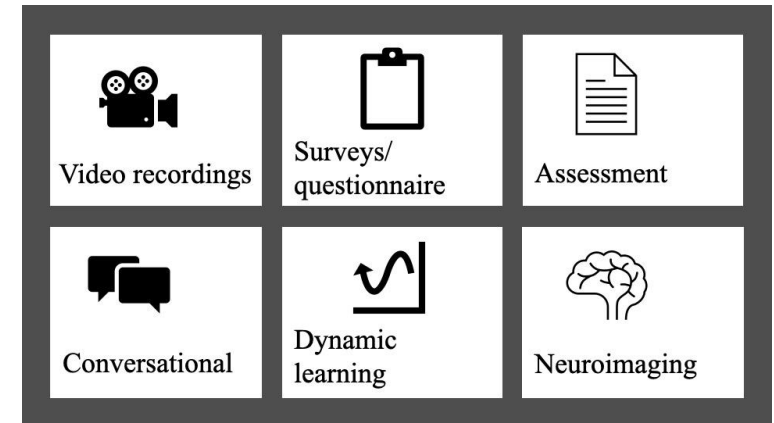
Challenge 1. Data privacy

Challenge 2. Flexibility across data types/users

Challenge 3. Scalability to large datasets

Challenge 4. Latency (in real-time)

Challenge 5. Noisy multi-modal data



# Challenges

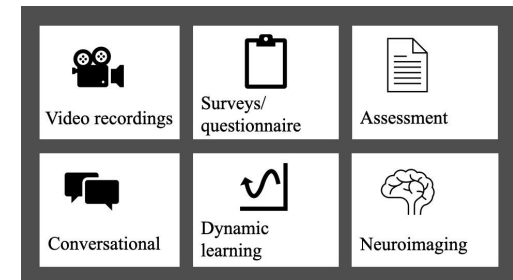
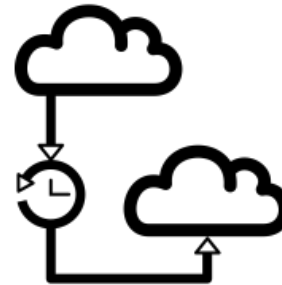
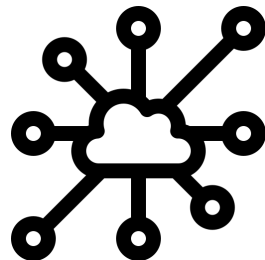
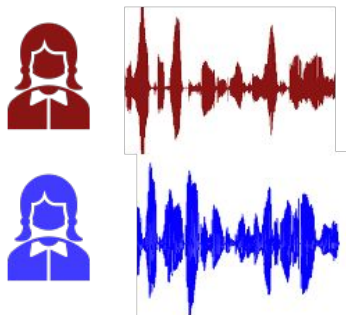
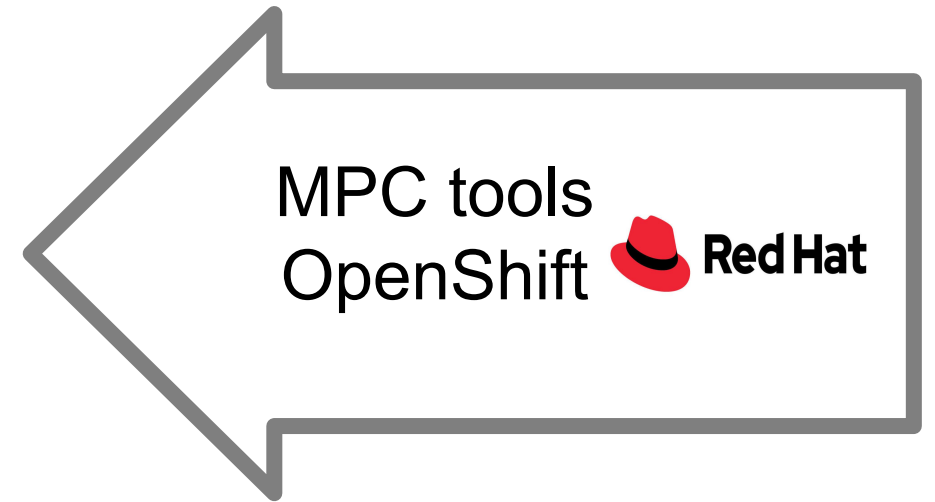
Challenge 1. Data privacy

Challenge 2. Flexibility across data types/users

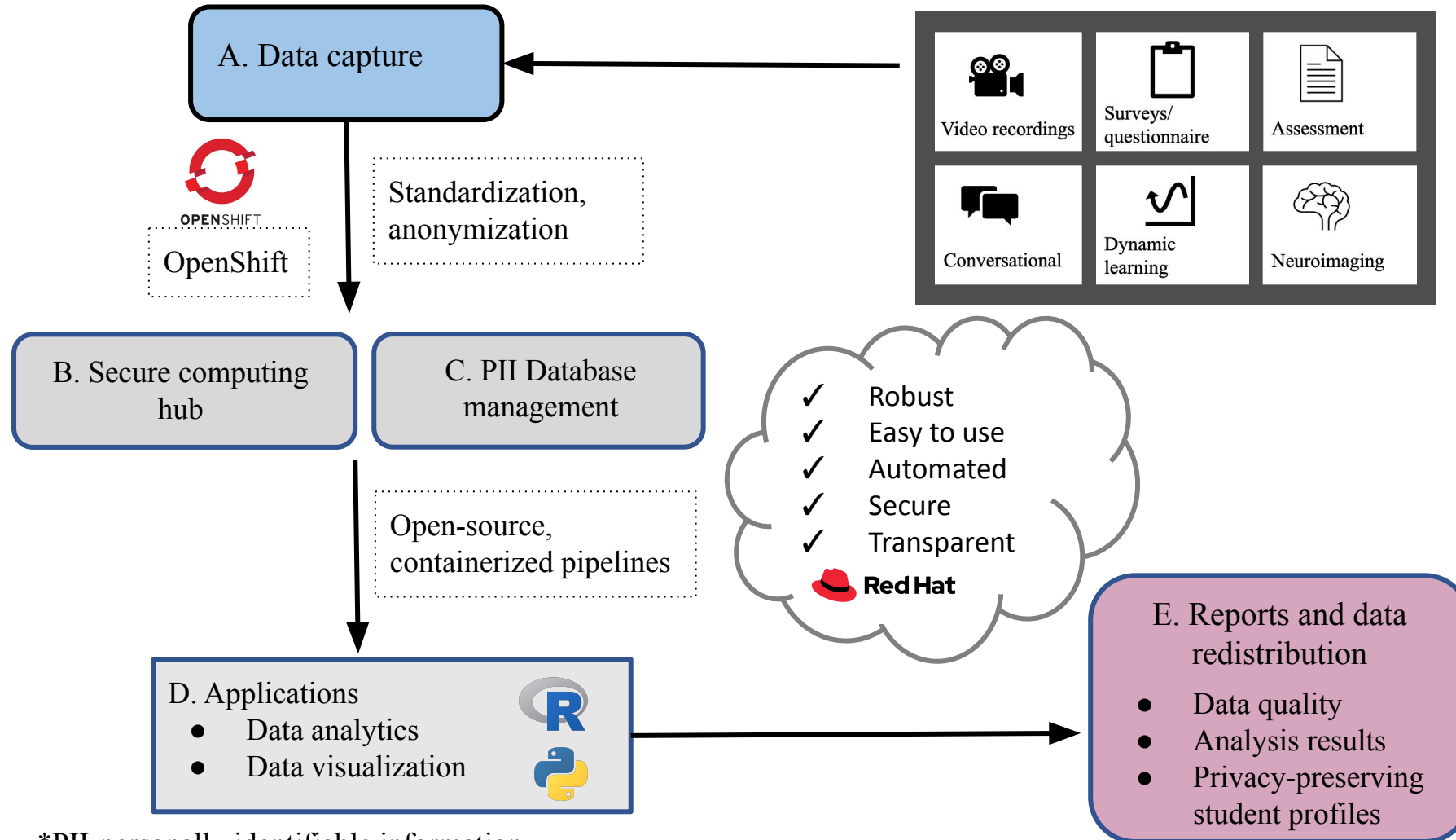
Challenge 3. Scalability to large datasets

Challenge 4. Latency (in real-time)

Challenge 5. Noisy multi-modal data

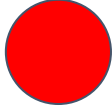


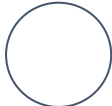



# Implementation of Repository



\*PII-personally identifiable information

# Progress and Next Steps

1. Curate large educational data 
2. Infrastructure for storage and validation of datasets 
3. Develop data processing and validation pipeline 
4. Develop APIs and OpenShift integrations 
5. Integrate Red Hat MPC 
6. Train and validate models for outcome predictions 