# dataspark

## THE UNIVERSITY OF RHODE ISLAND

## Prospectus

# SUSTAINING RHODE ISLAND'S STATEWIDE LONGITUDINAL DATA SYSTEM

# EXECUTIVE SUMMARY

## TODAY

DataSpark maintains and operates the State of Rhode Island's **statewide longitudinal data system** (RI SLDS) for the purposes of supporting the State's policy and decision-making priorities, namely **answering policy questions** and **supporting research** around education and workforce outcomes as well as **evaluating programs** and **measuring investments** in education and workforce development.

The RI SLDS is a State resource. As custodians of the RI SLDS, DataSpark serves **Rhode Islanders**: policymakers, community leaders, academic researchers, and the public at large.

## TOMORROW

DataSparks exists to foster the use of integrated, longitudinal data for research, evaluation and cross-sector problem solving. Our parallel requests of legislative **codification** and state **appropriations** will ensure the continued existence of DataSpark and the RI SLDS for the benefit of the State of Rhode Island.

# INTRODUCTION

## MISSION

DataSpark connects data and people to inform, empower and inspire innovative decision making.

## TEAM

Successfully operating a statewide longitudinal data system requires personnel skilled in data engineering, analysis, visualization, and governance. DataSpark currently has four employees, including data engineers to optimize the accuracy and efficiency of linkage and data analysts to conduct equity-driven analysis.

## LOCATION

Statewide longitudinal data systems are commonly located within executive agencies, or within public universities. In other states, these systems may also be stand-alone entities. DataSpark has been housed at a local nonprofit organization as well as within the State's public university.

# DATASPARK OF TODAY
## BENEFITS

DataSpark maintains and operates Rhode Island's **statewide longitudinal data system** (RI SLDS). The RI SLDS centralizes more than 50 datasets from eleven sources over three decades. It currently links data from **early childhood, through postsecondary education, and into the workforce**. For nearly 30 years, DataSpark has used the power of data to inform, empower, and inspire innovative decision making and research.

## BENEFITS

Fosters cross-sector problem solving

Drives accountability & investment decisions

Improves decision making & research

Supports federal & state reporting

Improves organizational capacity & efficiency

Enhances data quality & usability

Increases data transparency & accessibility

# DATASPARK OF TODAY
## FEATURES & USES

### DATA STEWARDSHIP

The RI SLDS centralizes more than 50 datasets from eleven sources over three decades. We currently link data from early childhood, through postsecondary education, and into the workforce. In addition to state data, the RI SLDS includes local, municipal data and public, federal data. DataSpark collaborates with contributing entities to improve quality and promote transparency. These durable partnerships, built over decades of cooperation, ensure accurate interpretation of and appropriate use of the data throughout its lifecycle.

### RESEARCH & POLICY

Using advanced statistical and data science methods, DataSpark analyzes cross-sector, longitudinal data. We provide additional capacity, supporting federal and state reporting. Similarly, DataSpark partners with nationally renowned university faculty, leveraging the RI SLDS, to conduct pioneering research, to evaluate state programs, and to improve the lives of Rhode Islanders. With approval from the data providers, we securely release anonymized RI SLDS data.

## DASHBOARDS & VISUALIZATIONS

DataSpark analyzes RI SLDS data, producing everything from traditional long-form reports to cutting-edge interactive dashboards. Whether a data story, infographic, or map, these visualizations distill complex information into insights. DataSpark analysts are not only expert users of the most common business intelligence tools, like Microsoft PowerBI and Tableau, but they also develop custom web-based applications to shed light on critical policy issues and promote evidence-based decision-making.

## ARTIFICIAL INTELLIGENCE (AI)

The RI SLDS links previously siloed data using a proprietary machine learning algorithm, custom-built by DataSpark engineers using the latest advances in AI. This algorithm, named MatchBot, ensures accurate person-level matches and reduces the need for time consuming human review of sensitive information. Over time, MatchBot learns, growing more confident in its decision making and further protecting data stored in the RI SLDS by decreasing the manual review necessary.

# DATASPARK OF TOMORROW
## PATH TO SUSTAINABILITY

To maximize benefits to the State, Rhode Island policymakers should pass legislation codifying RI SLDS and guaranteeing predictable funding. Together **codification** and **appropriations** will ensures that the RI SLDS is secure, responsive, and scalable to meet future State data needs.

This legislation should
- formalize the governance of the RI SLDS,
- require data sharing from education, workforce agencies, health, and human services agencies, and
- ensure reliable, sustainable funding.

Legislation formalizing governance structures builds on DataSpark's strong partnerships and ensures trust and transparency among stakeholders. Cross-agency governance, as exists for Connecticut's Preschool through 20 and Workforce Information Network (P20 WIN) and the Maryland Longitudinal Data System (MLDS), allow State leadership to prioritize research and policy questions.

The RI SLDS currently links data from early childhood through postsecondary education and into the workforce. A modern, robust system like the Kentucky Center for Statistics (KYSTATS), also integrates health and human services data as well as information about the environment, justice, and transportation. Requiring data sharing is the first step toward ensuring accurate, cross-sector, and timely reporting.
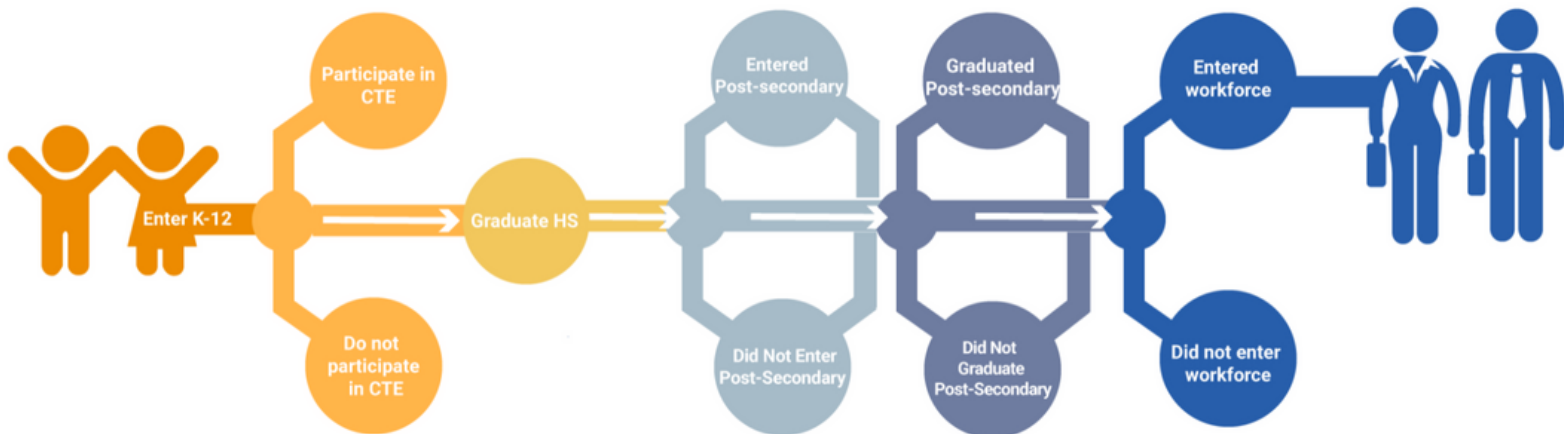
Finally, codification eliminates duplicative shadow data systems, providing a better return on state investment. Moving from ad hoc funding arrangements to appropriations increases the stability of the system and its overall benefit.

# DATASPARK OF TOMORROW
## PATH TO SUSTAINABILITY

These goals are only possible with **policymakers' support**. Leaders in the Governor's office, executive agencies, and the legislature are uniquely positioned to ensure the continued existence and expansion of the RI SLDS for the benefit of Rhode Islanders.

Once sustained, DataSpark can leverage to RI SLDS to provide additional products and services, beyond existing analytics and visualizations offerings. The RI SLDS can inform policymaking and support program evaluation related to career pathways, economic opportunities, healthcare access, housing stability, and more.

# DATASPARK OF TOMORROW
## PRODUCTS & SERVICES

**1**

### STATEWIDE

provides annual updates to the **Rhode Island Talent Dashboard** and develops a **COVID-19 Recovery Dashboard**; identifies **family & household connections;** produces a **unified Rhode Island address list**

**2**

### AGENCY-SPECIFIC

operates and maintains **agency-specific data warehouses**, providing real-time access to analysis-ready data; responds to **ad hoc requests** from agencies, the General Assembly, and the Governor's Office**;** increases capacity in **predictive analytics**

**3**

### THIRD-PARTY

serves as the **data releasing arm**, collaborating with each agency's designated research partners; builds a **data enclave**, a secured remote desktop environment for approved reseachers

# DATASPARK OF TOMORROW
## OPPORTUNITIES

As DataSpark integrates more health and human services data into the RI SLDS, it can better address equity challenges and better answer policy questions. By investing the RI SLDS, policymakers can drive wholistic improvements in wellbeing for all Rhode Islanders.

1. How does kindergarten readiness impact third grade reading proficiency?

2. How is the COVID-19 pandemic and distance learning affecting children's mental health?

3. How can Rhode Island improve the training and retention of high-quality early childhood educators?

4. What are the educational outcomes of youth involved with foster care or the justice system?

5. Which postsecondary students are food and housing insecure, and how can policymakers support these students?

6. How is the COVID-19 pandemic impacting employment opportunities, benefits eligiblity, housing stability, and healthcare access?

7. Which industry-recognized credentials offer its participants the highest return on investment?

8. Which Rhode Island communities provide the highly skilled workforce required in today's competitive global economy?

9. How do Rhode Islanders' commuting patterns vary by industry and wage?

10. How does adult education impact benefits utilization and civic participation?

11. Has the shift to remote work exacerbated Rhode Island's "brain drain", and which industries are most affected?