Secure, Standalone Data Intelligence for Regulated Industries

DATA WARP

Your Data. Your Machine. Your Insights.

The Enterprise Data Impasse

In the modern enterprise, the need for AI-driven insights is in direct conflict with the mandate for absolute data security. **Industry analysis shows over 80% of data science projects involving sensitive PII, PHI, or financial data are delayed or canceled due to security and compliance hurdles.**

Cloud-based analytics platforms introduce unacceptable data exposure risks, while legacy on-premise tools are too complex and slow for today's agile data teams. This creates an impasse: innovate with risk, or secure with stagnation. Until now.

DataWarp is a research-proven integrated, standalone platform that resolves this conflict, providing a high-performance data preparation and AI workbench that runs 100% within your own secure, air-gapped environment.

Key Capabilities

Zero-Data-Exposure AI

Leverage the power of large language models for insight and query generation with our proprietary, metadata-driven architecture. We guarantee no row-level data ever leaves your secure network, eliminating the primary barrier to AI adoption.

Unified Data Workbench

Achieve insights up to 80% faster by eliminating tool sprawl and context-switching. An integrated environment combines a high-performance data grid, advanced data cleaning and transformation tools, a powerful SQL editor, a guided formula engine, and an automated visualization suite.

Automated Pipeline Benchmarking

Move beyond guesswork. Quantitatively measure the impact of your data cleaning workflows by automatically benchmarking pipeline performance and the accuracy of downstream machine learning models. Prove the ROI of your data quality initiatives.

100% Standalone Architecture

Install DataWarp as a simple, self-contained desktop application on your Windows machine. With all data processing and analysis running 100% locally, you achieve ultimate data sovereignty. This architecture, which keeps your sensitive data off the cloud, simplifies compliance with GDPR, HIPAA, and other strict regulations.

Email: boruahpriyanuj2004@gmail.com
LinkedIn: www.linkedin.com/in/priyanuj-boruah

The DataWarp Advantage

Unmatched in the Industry

DataWarp was purpose-built to fill a critical gap left by other solutions. We are not a better cloud tool; we are the definitive alternative for when security cannot be a compromise.

vs. Public Cloud Platforms

Complete Data Sovereignty. Public clouds demand you compromise on security. We don't. DataWarp installs entirely on your machine, guaranteeing 100% data sovereignty and security that cloud vendors cannot match.

vs. Legacy ETL Tools

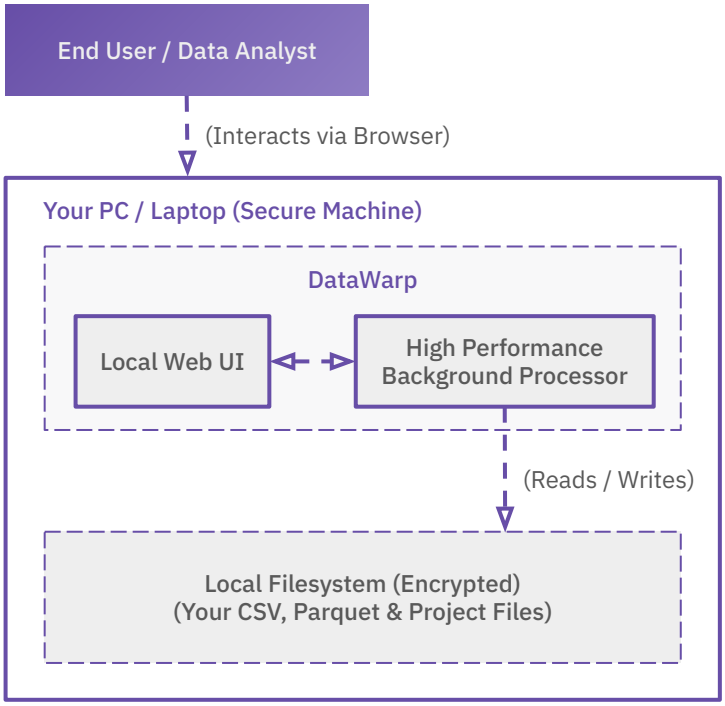
Agility and Speed. Legacy tools are slow, complex, and stifle innovation. DataWarp is built for speed, empowering your team to explore and transform data at the pace of modern analytics.

vs. Open-Source Frameworks

Governance and Accessibility. Code-only frameworks are powerful but create chaos—they lack governance, auditability, and are unusable by 90% of your team. DataWarp provides a governed platform that makes powerful data prep secure, standardized, and accessible to everyone.

DataWarp Secure Architecture

The platform is engineered for ultimate security. As a standalone desktop application, all processes run locally on your machine under your user account. This self-contained architecture eliminates the need for external services, open network ports, or complex configurations, ensuring the principle of least privilege. The entire system operates on your PC, with no inbound connections required. Outbound connections are limited to two specific, secure functions: a call to our secure identity provider for user authentication, and an optional, metadata-only API call for the AI features.



Data Workflows · Automated Analysis · Interactive Visualizations · User-Friendly Interface · Security-First Architecture · Zero-Exposure Advanced AI · Integrated SQL & Expressions · Export-Ready Pipelines

Email: boruahpriyanuj2004@gmail.com
LinkedIn: www.linkedin.com/in/priyanuj-boruah

Use Cases

Optimize AI & Machine Learning: Directly improve the accuracy and performance of your most critical models. Use DataWarp's integrated benchmarking to quantitatively prove which data preparation pipelines deliver the most predictive power, turning data quality into a measurable competitive advantage.

Accelerate Analytics & BI: Empower your data analysts and business users to move from raw data to insights in a fraction of the time. Replace slow, manual, and error-prone spreadsheet-based cleaning with a fast, intuitive, and governed platform that boosts team productivity.

Simplify Regulatory Compliance: Meet the strictest data residency and security requirements for GDPR, HIPAA, CCPA, and more. DataWarp's 100% on-premise architecture gives you a complete audit trail and guarantees you maintain absolute sovereignty over your sensitive data assets.

Operationalize Data Pipelines: Move beyond one-off scripts and ad-hoc cleaning. Design, save, and automate multi-step data cleaning workflows with the visual pipeline creator. Generate portable Python code to integrate your standardized processes into production MLOps or data engineering workflows.

Secure AI for Private Data: Unlock the power of large language models on your most sensitive datasets. Use DataWarp's revolutionary Zero-Data-Exposure AI to ask complex questions in natural language, getting answers in seconds without a single row of data ever leaving your secure network.

Unify and Standardize Disparate Data: Rapidly ingest, merge, and conform data from different business units and systems. Use DataWarp's powerful transformation engine to create a single, high-quality, analysis-ready dataset, eliminating data silos and establishing a trusted source of truth.

Technical Specifications

Component	Specification
Deployment Model	**Standalone Desktop Application delivered via a standard Windows Installer (.exe).**
Deployment Environment	**Windows desktops and laptops.**
Host Environment	**OS:** Windows 10 or Windows 11 (64-bit). **Resources:** Recommended 4 CPU Cores and 8 GB RAM.
Core Technology	**Engine:** Proprietary In-Memory Analytics. **UI:** Modern, Browser-Based Interface. **Data Storage:** Local Filesystem. **Authentication:** Secure user authentication managed via Supabase (SOC2 Type 2 Compliant).

Email: boruahpriyanuj2004@gmail.com
LinkedIn: www.linkedin.com/in/priyanuj-boruah

Research-Proven: The Three Pillars of DataWarp

Our innovation is transparent, published, and permanently archived at the Harvard Dataverse. We invite you to scrutinize the science that powers our platform.

Pillar 1: Security

Zero Data Exposure: A New Framework for Enabling Generative AI on Private Enterprise Data

This paper proves why common approaches to AI are fundamentally insecure for enterprise use. It establishes the rigorous, security-first framework that underpins DataWarp's **Zero-Data-Exposure AI**, the only architecture proven to be both analytically powerful and architecturally secure for sensitive data.

https://doi.org/10.7910/DVN/FZMD31

Pillar 2: Performance

SMART-IMPUTE: A Time-Efficient, ANN-Based Algorithm for Practical Imputation

This paper solves the critical speed vs. accuracy trade-off in data cleaning. We introduce and validate the Smart-Impute algorithm, proving it is up to **12.2x faster** than standard methods while retaining state-of-the-art accuracy, enabling agile workflows on enterprise-scale data.

https://doi.org/10.7910/DVN/TJ7WRT

Pillar 3: Strategy

The Laws of Anomaly: A Framework for Regression Model Selection

This landmark study transforms the art of model selection from an intuition-based heuristic into a formal science by establishing the **first data-driven laws of applied machine learning**. It proves the "silver bullet" role of robust models that win with **performance margins exceeding 1500%** in the face of hidden data anomalies.

https://doi.org/10.7910/DVN/VH9JJA

Validated by World-Leading Institutions

Our commitment to academic rigor is absolute. Our work is developed to the highest standards and validated by the world's leading institutions.

HARVARD Dataverse

All three foundational research papers are permanently archived and citable via the **Harvard Dataverse**—a world-leading research data repository from **Harvard University**.

IIT Madras

The founder and principal researcher behind this work is a top student at the **Indian Institute of Technology Madras (IIT Madras)**, recognized by the Indian government as the **#1 technical and engineering institute in the nation.**

About the Founder

The architect of this research and the DataWarp platform is **Priyanuj Boruah**. His work is built on a foundation of elite academic achievement, beginning with his distinction as the **2022 CUET-UG Computer Science National Topper**. As an independent researcher and student at IIT Madras, he authored the "Laws of Anomaly" and the foundational papers on secure AI and high-performance imputation that serve as the pillars of DataWarp.

Holding a unique and powerful **Law in Data Science**—the first of its kind in the world—Priyanuj has engineered DataWarp with a rare, integrated understanding of both technical possibility and the legal and regulatory realities of enterprise data. This is why our commitment to security and privacy is not just a feature; it is **embedded in our architectural DNA**.

Email: boruahpriyanuj2004@gmail.com
LinkedIn: www.linkedin.com/in/priyanuj-boruah