Ontara: Toward a Computable Semantic Infrastructure for K–16 Mathematics

Executive Summary

Across the United States, mathematics instruction operates without a unified, computable representation of what mathematical knowledge *is*, how it develops, and how it can be reliably measured. Standards such as the Common Core State Standards (CCSS) and Achieve the Core (ATC) provide textual descriptions of learning goals and broad learning progressions, but they do not supply the semantic structure required for advanced assessment, personalized learning, or AI-driven instructional systems.

Ontara refers to a proposed ontological framework designed to supply this missing semantic layer. Its purpose is not to replace existing standards, but to provide the rigorous, machine-readable conceptual structure needed to transform teaching, assessment, and learning. The scope of this effort is extraordinarily broad: it requires formalizing K–16 mathematics in a way that is mathematically coherent, pedagogically sound, computationally tractable, and sufficiently expressive to support reasoning across millions of students and a wide diversity of problem types.

This white paper outlines the **problem space**, the **high-level capabilities** required to address it, the **core technical challenges**, and the **potential impact on learners and the education market**.

1. The Problem: A Fragmented, Informal, and Incomputable Knowledge Landscape

1.1 Standards Without Structure

Mathematics standards are expressed as natural-language statements describing skills, concepts, and expectations. These documents lack:

- **Formal mathematical specificity**
 (e.g., they rarely distinguish between *function* as a relation, as an input–output rule, or as a modeling construct).
- **Explicit semantic structure**
 (e.g., how factoring relates to solving quadratics; how ratios progress into slope).
- **Machine-actionable coherence**
 (e.g., represented prerequisite relationships, cluster hierarchies, or cross-domain dependencies).

Although the standards community has made meaningful progress in organizing content—domains, clusters, and learning progressions—this organization is **not computable**, **not rigorous enough for AI systems**, and **not expressive enough to encode mathematical meaning.**

1.2 Assessment and Instructional Systems Built on Flat Tags

Current instructional and assessment systems generally treat standards as atomic labels (“A-REI.B.3”, “8.F.A.3”). This leads to:

- Assessment misalignment and inconsistent rigor
- Poor identification of conceptual vs. procedural understanding
- Limited personalization
- Inability to detect domain interactions (e.g., proportional reasoning underlying linearity)

Without structural meaning attached to these labels, systems cannot understand *what* a student knows—only *how* they performed on isolated items.

1.3 Absence of a Coherent Student Mastery Model

Even advanced adaptive platforms rely heavily on item-level statistics rather than conceptual representations of mathematical knowledge. This produces:

- Unreliable cross-grade predictions
- Difficulty detecting “slip zones” where procedural skills mask conceptual weakness
- Weak models of time-to-mastery and learning trajectories
- Inability to generalize support from one domain to another

Without an underlying ontology, mastery modeling remains shallow and brittle.

2. The Ontara Vision: A Semantic Infrastructure for Mathematics

Solving the underlying problem requires a **fundamental re-engineering of how mathematics knowledge is represented, organized, and computed upon**. Ontara encompasses a multi-layered architecture that would enable this shift.

2.1 A Computable Knowledge Graph of Mathematics Standards

A semantic graph is required to capture:

- **Nodes** representing standards, sub-skills, and conceptual kernels
- **Edges** encoding prerequisite relationships, conceptual dependencies, and structural links
- **Hierarchies** representing domains, clusters, and learning progressions
- **Crosswalks** across grade levels and other frameworks

This graph must unify structural representations (e.g., algebraic properties), pedagogical representations (e.g., problem types), and curricular sequences.

2.2 A Mathematical Ontology Layer

Underlying all standards is actual mathematics: functions, operations, relations, symbolic structures, representations, and invariants. A complete system requires:

- Formal definitions of mathematical objects (expressions, equations, ratios, data sets)
- Relationships between them (is-a, part-of, generalizes, applied-to)
- Mappings from natural-language descriptions to mathematical constructs
- Representations that can support both symbolic manipulation and conceptual reasoning

This ontology must be compatible with but not constrained by existing systems such as OntoMathPRO; it must incorporate school-level constructs that do not exist in university-level mathematical ontologies.

2.3 Problem Archetype Engine

To reason about assessment, the system must model:

- Canonical problem families (e.g., “linear equations with variables on both sides”)
- Common representations (graphs, tables, symbolic expressions, verbal contexts)
- Layers of abstraction and rigor
- Structural transformations and equivalences
- Common misconceptions and error types

A problem archetype engine provides the generative backbone for:

- Assessment creation
- Thin-slicing and scaffolding
- Personalized re-teaching
- Skill diagnostics
- Rigor balancing

2.4 Student Mastery Graph and Learning Trajectory Modeling

A deep conceptual model requires:

- Probabilistic mastery states for each conceptual node
- Trajectory models capturing growth, decay, and learning rate
- Cross-domain interactions (e.g., how fraction fluency affects algebra readiness)
- Evidence integration from multiple item types and modalities
- Longitudinal coherence across years

This model must be able to scale to **thousands of students**, **millions of interactions**, and **multiple domains**, while maintaining interpretability for teachers and administrators.

3. Why This is an Exceptionally Hard Problem

The Ontara vision represents a **grand challenge** at the intersection of mathematics, linguistics, cognitive science, computer science, and psychometrics. Several dimensions of difficulty stand out.

3.1 Formalizing Natural-Language Standards into Mathematical Semantics

Educational standards are not written in precise mathematical language. Translating them into formal semantics requires:

- Mathematical interpretation
- Semantic parsing
- Disambiguation of cognitive tasks
- Hierarchical structuring
- Integration of pedagogical meaning

This alone constitutes a significant research undertaking.

3.2 Bridging Mathematical Knowledge and Student Cognition

Formal mathematics and K–16 learning progressions do not align neatly. Mathematics has a single coherent structure; learning trajectories do not. Building a system that simultaneously respects:

- Mathematical truth
- Cognitive development
- Instructional sequences
- Standards framings

is profoundly challenging.

3.3 Representing Problems in a Canonical Mathematical Form

Mathematical problems appear as:

- Symbolic expressions
- Word problems
- Graphs
- Tables
- Diagrams

Creating a single canonical representation capable of capturing their structure and meaning is complex, particularly when considering:

- Equivalent forms
- Implicit context
- Multi-step reasoning
- Modeling tasks
- Varying levels of abstraction

3.4 Scaling a Knowledge Graph with High Fidelity

The required ontology spans:

- 500+ standards
- Thousands of conceptual sub-nodes
- Tens of thousands of edges
- Millions of potential problem variants
- Continuous updates across years and curricula

Maintaining consistency, preventing contradictions, and enabling fast computation is non-trivial.

3.5 Integrating Symbolic Mathematics, Natural Language, and Pedagogical Intent

Few systems can operate simultaneously across:

- Symbolic algebra
- Graph reasoning
- Natural-language understanding
- Cognitive modeling
- Psychometric evaluation

This integration is extraordinarily difficult and lies at the frontier of AI research.

3.6 Creating Reliable and Valid Assessment Signals

Each problem must be aligned not only to content but also to *rigor*, *misconception profile*, *construct coverage*, and *psychometric properties*. This requires:

- Consistency
- Reliability
- Validity
- Bias detection

across problem archetypes—another significant challenge.

4. Expected Educational Impact

A system capable of addressing the challenges above would deliver transformative outcomes.

4.1 Personalized Learning with Conceptual Integrity

Students would receive learning experiences tailored to:

- Their conceptual profile
- Their misconceptions
- Their learning pace
- Their trajectory toward college and career readiness

This personalization would be principled, not ad hoc.

4.2 Coherent Progression Through Mathematical Ideas

Students would experience mathematics as a connected whole rather than a series of disconnected skills. The system would make visible:

- How topics connect
- Why they matter
- What comes next
- How earlier ideas support later ones

This supports both deep understanding and persistence.

4.3 High-Quality, Reliable Assessment at Scale

Districts and educators could generate assessments that are:

- Balanced in rigor
- Psychometrically robust
- Aligned to coherent learning trajectories
- Sensitive to conceptual vs. procedural differences
- Easily convertible into personalized re-assessments

4.4 Teacher Empowerment

Teachers gain:

- Actionable insights into student learning
- Automatically generated thin-slicing sequences
  High-quality item banks
- Visibility into misconceptions
- Reduced planning and assessment burden

This shifts teacher time toward instruction and relationships rather than paperwork.

5. Market Opportunity and Strategic Positioning

5.1 Core K–12 Mathematics Market

Mathematics is the most heavily assessed subject in the United States and the subject with the greatest need for remediation. A semantic infrastructure would serve:

- State agencies
- Districts
- Curriculum publishers
- Assessment companies
- Tutoring providers
- Digital learning platforms

5.2 Next-Generation AI Education Systems

All generative AI–based education tools will require a grounding ontology that prevents hallucination and ensures instructional soundness. Ontara provides exactly this foundation.

5.3 Higher Education and Workforce Development

Math pathways in community colleges and general education mathematics suffer the same alignment problems as K–12. A coherent ontology is valuable for:

- Placement
- Remediation
- Pathway acceleration
- Adult learning
- Workforce upskilling

5.4 Global Opportunity

Many countries face similar challenges in mathematics education coherence. A scalable semantic ontology offers international expansion potential with minimal localization effort.

6. Conclusion

The need for a computable semantic infrastructure for mathematics is clear. Without it, personalized learning remains shallow, assessments remain inconsistent, and AI-based instructional systems lack the grounding required for reliability and rigor.

Developing such an ontology—Ontara—represents a **major scientific, engineering, and educational challenge**. It requires integrating mathematical structure, cognitive development, curriculum design, psychometrics, symbolic reasoning, and large-scale data systems into a single coherent architecture.

The scope is vast.
The problem is hard.
But the potential impact—on students, teachers, districts, and the global education market—is equally enormous.

Ontara aims to supply the missing backbone that modern mathematics education requires.