Black Sea Journal of Engineering and Science
doi: 10.34248/bsengineering.1991498

Open Access Journal
e-ISSN: 2619 – 8991

Research Article
Volume 9 - Issue 5: 2393-2403 / September 2026

DEVELOPMENT AND WEB-BASED VISUALIZATION OF A DOMAIN-SPECIFIC SEMANTIC MODEL FOR COFFEE CULTURE AND PRODUCTION PROCESSES

Ömer DOĞAN[1], Alpay DORUK[2]*

[1]*Bandirma Onyedi Eylul University, Faculty of Engineering and Natural Sciences, Department of Computer Engineering, 10200, Balikesir, Türkiye*
[2]*Bandirma Onyedi Eylul University, Faculty of Engineering and Natural Sciences, Department of Computer Engineering, 10200, Balikesir, Türkiye*

Abstract: The global agricultural supply chain is a complex network with different data silos, complicated cultivation parameters and highly subjective sensory evaluation metrics. Formal knowledge representation is a must, if real semantic interoperability is to be achieved across this multidimensional domain. This work consists of the rigorous design, implementation and quantitative evaluation of the "Coffee Culture and Production Processes Ontology" and a custom client-side web-application named the "Coffee Ontology Explorer". The semantic model, developed with standard ontology design principles and the Web Ontology Language (OWL), captures the entire coffee lifecycle from botanical taxonomy and harvesting methods to specific roasting profiles and subjective brewing evaluations. Meanwhile, the React.js based ontology explorer tackles the ongoing limitations of traditional, heavy-client ontology visualization tools by utilizing a browser-native document parsing application programming interface for high efficiency, serverless ontology parsing. The quantitative estimation of the developed semantic model gives a 0.297 relation richness indicator, which confirms a highly interconnected and structurally complex knowledge graph, and not a simple taxonomic flat. Moreover, the performance profiling of the web application shows less than one second for parsing and less than 200 milliseconds for rendering complex graphical data even for large ontological hierarchies. This technical framework is a highly scalable, multilingual semantic infrastructure that is uniquely suited for integration into larger industrial traceability systems and provides a fundamental template for computer scientists, agritech developers, and agricultural stakeholders who want to exploit semantic web technologies for precision agriculture.

Keywords: Semantic web, Knowledge graph, Description logic, Web application, Coffee production, Supply chain traceability

***Corresponding author:** Bandirma Onyedi Eylul University, Faculty of Engineering and Natural Sciences, Department of Computer Engineering, 10200, Balıkesir, Türkiye
E mail: adoruk@bandirma.edu.tr (A. DORUK)
Ömer DOĞAN https://orcid.org/0000-0003-4815-8619
Alpay DORUK https://orcid.org/0000-0002-6190-288X

Received: July 10, 2026
Accepted: August 10, 2026
Published: September 15, 2026

Cite as: Doğan, Ö., & Doruk, A. (2026). Development and web-based visualization of a domain-specific semantic model for coffee culture and production processes. *Black Sea Journal of Engineering and Science,* 9(5), 2393-2403.

1. Introduction

The modern agriculture sector is becoming increasingly dependent on advanced data-driven paradigms that optimize physical production, ensure end-to-end supply chain traceability, and comply with strict environmental sustainability standards (Goldstein et al., 2021; Smaili and Kabbaj, 2025). In this bigger picture, the global coffee industry is more vulnerable. Coffee is not merely a bulk agricultural commodity but a highly nuanced and chemically complex product, in which small variations in altitude of cultivation, post-harvest processing methods and thermal roasting temperatures can have a profound effect on the final chemical composition and subjective sensory profile of the beverage (Illy and Viani, 2005).

The adoption of Internet of Things sensors and distributed ledger technologies for improved traceability of coffee supply chains is expanding quickly (Kamble et al., 2020; Lin et al., 2020; Ramos et al., 2023), but the data structures that support these innovations are still very fragmented. Agricultural and industrial terminology is very context specific and geopolitically bounded making it very difficult to integrate, query and analyze data efficiently across various digital agencies and international stakeholders (Joo et al., 2016).

Semantic Web Technologies provide a very strong, mathematically formalized framework (Gruber, 1993) for dealing with this fundamental heterogeneity. Ontologies are defined as explicit formal specifications of a shared conceptualization, so that machine agents can interpret data based on deep contextual awareness rather than simple syntactic pattern matching (Dooley et al., 2018; Guarino et al., 2009). Ontologies impose tight classes, object properties and description logic restrictions and transform independent and flat data points into rich and interconnected semantic knowledge graphs (Keet, 2019; Noy and McGuinness, 2001). However, the rigorous development of a domain specific ontology has to be complemented by accessible user centric visualization

and exploration tools to be of any practical utility. In the past, the use of ontological data was limited to either dedicated desktop applications or heavy server-side processing environments, thus creating prohibitively high barriers to entry for non-technical domain experts such as agronomists, coffee roasters, and supply chain logistics managers (Horridge, 2007; W3C OWL Working Group, 2012).

The use of semantic web technologies in the agricultural domain has gained substantial empirical traction, mainly driven by the growing worldwide demand for precision agriculture and transparent supply chain logistics (Goldstein et al, 2021; Smaili and Kabbaj, 2025). Important groundwork has been laid by prominent macro-ontologies and controlled vocabularies such as AGROVOC from the Food and Agriculture Organization (Liang et al., 2006) and the widely used FoodOn framework (Dooley et al., 2018) on which agriculture-specific vocabularies have been built. As a consortium-driven project, FoodOn aims to develop a complete farm-to-fork ontology to accurately and consistently describe foods across different global cultures (Dooley et al., 2018). FoodOn does offer some basic taxonomic representations of coffee, but its scope is necessarily broad. It typically views coffee only as a source of plant food or extracted beverage, and lacks the fine-grained, domain-specific axioms required to mathematically model the complex causal dependencies between coffee botanical varieties, high altitude cultivation stressors, specific thermal roasting curves, and final sensory cupping profiles (Angelica and Ferdinand, 2017).

The emerging computational literature stresses the need for specialized micro-ontologies to extend these foundational macro-models into highly specialized verticals. Much of the research on the coffee supply chain has been focused on traceability, using distributed ledger technologies and intelligent contracts to trace beans from farm to cup (Mauladi et al., 2022; Nuraisyah et al., 2025; Ramos et al., 2019). However, these advanced implementations of blockchain are often plagued by the classic “garbage in, garbage out” paradigm. Without the semantic alignment underneath, the raw data recorded on the ledger is functionally opaque to machine reasoning algorithms (Kamble et al., 2020; Ramos et al., 2019). Moreover, there have been existing expert systems based on mathematical analysis of Arabica coffee beans considering aroma and flavor inputs, with preliminary OWL models, but these systems generally lack complete and two-way structural linkages back to specific harvesting and processing variables (Goldstein et al., 2021). Therefore, a dedicated highly granular semantic model for the coffee industry still represents a critical gap in the existing literature.

The visualization of complex ontological data in an interactive way is a well-documented computational challenge in the wider Semantic Web community. Traditional knowledge engineering tools, and especially the Protégé integrated development environment, provide advanced editing and reasoning capabilities, but they are built on highly complex user interfaces and are strictly bound to desktop Java Virtual Machine environments (Keet, 2019; Horridge, 2007; W3C OWL Working Group, 2012). To democratize the access to semantic data and to promote a wider adoption across the industry, web-based visualizers have been extensively researched and proposed. Tools such as WebVOWL translate raw OWL syntax into visual force-directed graphs (Bach et al., 2011; Lohmann, et al., 2014). Other tools depend on heavy server-side rendering pipelines that take an uploaded ontology file, process it with backend Java frameworks, and return basic visual coordinates to the thin client (Dudáš et al., 2018; Katifori et al., 2007).

However, using server-side parsing incurs large network latency, requires maintenance of persistent and expensive back-end infrastructure, and poses critical data privacy issues for corporations using proprietary industrial ontologies that cannot be transmitted over public networks (Dudáš et al., 2018). In parallel, the rapid evolution of state-of-the-art JavaScript frameworks, such as React.js, and extremely optimized visualization of data libraries like D3.js and Recharts, have paved the way for a paradigm shift towards pure and unadulterated client-side processing (Coene, 2018; Chart.js documentation, 2026). React’s component-based, functional architecture allows for the declarative rendering of complex node-link diagrams and deeply hierarchical data, dynamically responding to changes in user state without needing to continuously hit the server (Lavanya et al., 2023; Leung and Salga, 2010). However, this technological change has not been accompanied by a rigorous analysis of the computational performance of native document object model parsing of OWL/XML documents strictly within a React state-management lifecycle in the peer-reviewed literature. This is a significant research gap that this technical study seeks to rigorously evaluate and fill.

To objectively measure the structural integrity, design quality and semantic richness of any newly developed ontology, quantitative mathematical metrics have to be used (Maynard et al., 2006). Qualitative expert evaluation or anecdotal validation alone is insufficient to establish the ultimate suitability of an ontology for automated machine reasoning and large-scale data integration. Some of the popular ontology engineering frameworks (Burton-Jones et al., 2005; Maedche and Staab, 2002; Tartir et al., 2005) define several important schema-level metrics, including Relationship Richness, Attribute Richness and Inheritance Richness.

These metrics provide detailed and quantitative information on the underlying design principles of the ontology, especially telling us whether the model is a simple flat hierarchical taxonomy or a complex highly interconnected knowledge graph that is capable of fully supporting sophisticated deductive reasoning algorithms (Keet, 2019; Tartir et al., 2005). For example, an ontology with very low relationship richness, built mainly around

simple parent-child "is-a" relationships, versus an ontology with high relationship richness, defining complex, real-world interactions between otherwise unrelated domain entities. These very formal mathematical definitions are part of the evaluation phase of this broad research and are used to rigorously validate the engineered coffee ontology.

This research focuses on these two technical challenges directly with a two-dimensional contribution. It first describes the precise engineering of a specialized domain ontology explicitly designed to map the intricacies of coffee culture and production processes. Second, a new, highly optimized web-based software architecture is presented, the Coffee Ontology Explorer, which uses modern JavaScript frameworks to load, query, simulate and visualize complex OWL files entirely within the client-side environment. This work combines formal knowledge representation with a very high performance, responsive web interface to provide an end-to-end semantic solution for the complex needs of the coffee industry.

2. Materials and Methods

The semantic model was formally developed with the architecture to be a deeply semantically rich knowledge graph for the entire physical and chemical lifecycle of coffee. The development lifecycle followed very closely the accepted ontological engineering methodologies (Noy and McGuinness, 2001) which specify a very iterative lifecycle of initial domain analysis, integration of existing ontologies, extraction of expert terminology, rigorous definition of the class hierarchy, mapping of object properties and formulation of logical constraints. The ontology was built using the Web Ontology Language (OWL) within the Protégé computational environment and strictly adhering to the standards of the World Wide Web Consortium (Angelica and Ferdinand,2017).

The primary existential entities in the coffee domain constitute the taxonomic backbone of the ontology. This gave rise to a deep and strongly structured hierarchical tree, which explicitly separated botanical living entities, physical states of the agricultural product, highly mechanized industrial processes, and subjective human sensory evaluations. The main ontological super-classes are defined with their respective hierarchical descendants as shown in Table 1.

Table 1. Complete hierarchical classes of the coffee culture and production processes ontology

Primary Super-Class (8)	Explicit Hierarchical Sub-Classes (19)	Total
CoffeePlant	CoffeeSpecies, CoffeeVariety	3
CoffeeBean	GreenBean, RoastedBean, GroundCoffee, EspressoSuitableCoffee (Inferred)	5
CoffeeProcessing	HarvestingMethod, ProcessingMethod, DryingMethod	4
CoffeeRoasting	RoastLevel, LightRoast, MediumRoast, DarkRoast	4
BrewingMethod	EspressoBrewing, FilterBrewing, ImmersionBrewing	4
TastingProfile	Flavor, Acidity, Body, Sweetness	5
Origin	Country, Region, Farm	4

The first thing to classify is the botanical origin. This is divided into specific taxonomic species (ex, Arabica and Robusta) and then into exact genetic cultivars and varieties (Bourbon or Typica). The main bean class represents the physical agricultural product itself. The main class is strictly categorized by the current state of the physical lifecycle, i.e. raw green beans, thermally processed roasted beans and mechanically pulverized ground coffee.

The industrial and agricultural activities are confined to specific processing classes. These are the agricultural extraction methods during harvesting, the post-harvest fermentation and washing processes, and the key drying techniques to achieve optimal moisture levels. Thermal processing is modeled as a separate class hierarchy, with detailed parameters for roasting levels from light to dark. The final human interaction with the product is modelled as specific classes of brewing method, modelling aqueous extraction methods such as espresso or immersion brewing. In an effort to incorporate the highly subjective nature of human taste, a complex sensory analysis class is developed that diverges into specialized sensory attributes such as specific flavor notes, perceived acidity, physical mouthfeel or body, and inherent sweetness. Finally, the geographical and geopolitical provenance is hierarchically organized from the macro-level of the producing country to specific regional zones and to the micro-level of the exact producing farm.

A full range of formal semantic properties was defined to transform the static, tree-like taxonomy into a highly relational, multidimensional knowledge graph. The Web Ontology Language explicitly distinguishes between Object Properties that express directed relationships between particular instances of two different classes, and Datatype Properties that relate a particular instance to a standardized literal value defined by XML Schema Datatypes (W3C OWL Working Group, 2012).

The object properties determine the complex logistical and chemical flow of the coffee bean in its entire life cycle. Strict mapping of domain constraints and range constraints guarantees rigorous semantic validation in automated reasoning tasks, ensuring that no illogical assertions are made in the knowledge base. The object properties implemented in the Coffee Culture and Production Processes Ontology are shown in Table 2.

The datatype properties are those that capture the exact quantitative variables and qualitative descriptors that are absolutely necessary for industrial quality control, sensory grading and precision agriculture (Goldstein et al., 2021; Gruber, 1993). These properties strictly utilize standard schema types to ensure computational interoperability (Table 3).

Table 2. Object properties implemented in the coffee culture and production processes ontology

Formal Property Identifier	Assigned Domain Class	Assigned Range Class	Semantic Role and Description
growsIn	CoffeePlant	Region	Associates a specific botanical plant instance with its geographic cultivation zone.
harvestedBy	CoffeeBean	HarvestingMethod	Defines the physical agricultural extraction method applied to the plant.
processedBy	CoffeeBean	ProcessingMethod	Links the raw agricultural cherry to its specific fermentation or washing process.
roastedTo	RoastedBean	RoastLevel	Defines the precise thermal transformation state achieved by the roaster.
hasOrigin	CoffeeBean	Origin	Links the final consumable product back to its geopolitical and physical provenance.
hasSpecies	CoffeeBean	CoffeeSpecies	Defines the exact botanical taxonomy of the specific harvested batch.
hasVariety	CoffeeBean	CoffeeVariety	Defines the specific genetic cultivar of the plant.
brewedWith	CoffeeBean	BrewingMethod	Maps a processed bean to its actual, real-world preparation method.
hasTastingProfile	CoffeeBean	TastingProfile	Connects the physical, chemical bean to its subjective human sensory evaluation.
producedIn	CoffeeBean	Farm	Provides micro-level logistical traceability down to the exact coordinates of the farm.
suitableFor	CoffeeBean	BrewingMethod	Serves as an expert recommendation link for optimal aqueous extraction.

Table 3. Datatype properties implemented in the coffee culture and production processes ontology

Formal Property Identifier	Assigned Domain Class	Assigned Range Class	Semantic Role and Description
Formal Property Identifier	Assigned Domain Class	XML Schema Datatype	Semantic Role and Description
altitudeGrown	CoffeePlant	xsd:integer	Cultivation altitude measured in precise meters above global sea level.
harvestYear	CoffeeBean	xsd:gYear	The standardized temporal stamp of the agricultural harvest.
roastDate	RoastedBean	xsd:date	The standardized temporal stamp of the industrial thermal processing.
acidityLevel	TastingProfile	xsd:integer	Standardized sensory scale rating ranging from 1 to 10 for perceived acidity.
bodyLevel	TastingProfile	xsd:integer	Standardized sensory scale rating ranging from 1 to 10 for physical mouthfeel.
sweetnessLevel	TastingProfile	xsd:integer	Standardized sensory scale rating ranging from 1 to 10 for inherent sweetness.
flavorNotes	TastingProfile	xsd:string	Unstructured text field for highly complex aroma notes, such as floral or citrus.
optimalBrewTemp	BrewingMethod	xsd:float	Target thermodynamic water temperature measured in degrees Celsius.
grindSize	BrewingMethod	xsd:string	Required physical particulate size categorization for proper extraction.
brewingTime	BrewingMethod	xsd:duration	Target temporal duration for optimal aqueous extraction of soluble compounds.

The intrinsic power of Description Logic ontology is its real computational power of automated, deductive reasoning (Guarino et al., 2009; Lohmann et al., 2014). By applying strict logical constraints to the defined classes using advanced mathematical set theory, the ontology enables automated semantic reasoners to infer entirely new classifications and relationships not explicitly asserted in the raw input data (Keet, 2019; W3C OWL Working Group, 2012).

The semantic model has a number of very complex logical constraints to represent real industry norms. The concept of coffee suitable for espresso extraction is presented as a precise mathematical intersection of a base class and an existential restriction. In particular, the reasoning engine automatically assigns any roasted bean with a known relationship to a roasting level defined as medium or dark as being suitable for espresso. This particular axiom enables an automated recommendation engine to classify any new agricultural batch introduced into the system dynamically without human intervention or manual reclassification.

Also, the notion of specialty coffee in the highly regulated industry is formalized through strict datatype property restrictions. The system mathematically encodes industry standard grading rules such as the rule that “any coffee bean grown at an altitude strictly greater than 1200 meters is classified within the specialty tier”. The system encodes key agronomic facts into machine-readable logic using datatype property restrictions. Beans grown at higher altitude are subjected to more severe environmental stress leading to more dense cellular architectures with slower maturation hence much more complex chemical acidity profiles when roasted (Illy and Viani, 2005).

Geographic inferences are also heavily encoded, so a coffee bean that has an origin relationship to the nation of Ethiopia is automatically tagged with a particular geographic tag. Additionally, we impose stringent cardinality constraints to ensure the integrity of the data as a whole. This means that each coffee bean can have only one relationship describing the botanical species of the coffee bean. The mathematics of this constraint ensures that the data is absolutely correct. There can be no taxonomic conflicts in a batch of products. This is a common mistake in badly designed relational databases. Finally, the mathematical bounding requires that a given brewing method must have a best brewing temperature greater than or equal to 85.0° Celsius, imposing thermodynamic preparation standards. This constraint is a quality control tool that automatically verifies that the

user-input recipes strictly follow the basic principles of thermodynamics for optimal extraction of coffee compounds. A wide range of query structures were carefully built so as to empirically demonstrate the functionality and industrial usefulness of the proposed semantic model using the SPARQL Protocol and RDF Query Language (W3C SPARQL Working Group, 2013). These fine-tuned queries show a model's inherent capability to navigate deep, complex graph relationships and expose multidimensional, useful information for supply chain stakeholders. In Table 4 mapping of competency Questions (CQs) to SPARQL Query objectives is shown.

Table 4. Mapping of Competency Questions (CQs) to SPARQL Query Objectives

CQ ID	Competency Question (CQ)	Target Semantic Path & SPARQL Objective
CQ1	Which coffee species/cultivars can thrive under high-altitude conditions (>1500m)?	Traverses CoffeePlant, CoffeeSpecies, and filters datatype altitudeGrown > 1500.
CQ2	Which coffee lots are inferred as suitable for espresso brewing without manual tag checks?	Queries the dynamically inferred class EspressoSuitableCoffee.
CQ3	What are the sensory profiles (acidity, body, flavor) of coffee batches originating from Ethiopia?	Traverses Origin (Ethiopia) → CoffeeBean → TastingProfile.
CQ4	Which coffee batches match specific subjective flavor descriptors (e.g., "floral")?	Performs regex/pattern filter on flavorNotes datatype property.

The first query structure is high-altitude agronomic filtering. This search involves a complex traversal between the physical bean, its strict botanical species taxonomy and its exact cultivation datatype in an effort to identify some botanical species that are doing well under extreme altitudinal stress. The execution path filters the huge knowledge graph to return only agricultural instances strictly belonging to the Arabica species, but also checks the integer value of the cultivation altitude to make sure that it exceeds the strict threshold of 1,500 meters. This question will be very useful to agronomists, who are interested in studying the resistance of some genetic cultivars to environmental stressors at high altitudes.

The SPARQL Syntax for CQ1 is shown below:

```
PREFIX coffee:
<http://www.semanticweb.org/ontologies/coffee#>
PREFIX xsd:
<http://www.w3.org/2001/XMLSchema#>
SELECT ?plant ?species ?altitude
WHERE {
 ?plant a coffee:CoffeePlant ;
      coffee:hasSpecies ?species ;
      coffee:altitudeGrown ?altitude .
 FILTER (?altitude > 1500)
}
```

The secondary query structure demonstrates the great power of inferential recommendation retrieval. Instead, the search directly targets the mathematically inferred class of coffee suitable for espresso, without having to search through the whole range of thermal roasting parameters manually. This approach to executing the query pushes the entire classification burden directly onto the underlying Description Logic reasoning engine. The query does not need to explicitly test or validate roast levels, as the semantic engine has already processed the raw data and dynamically inferred the exact set of eligibles, based on the foundational axioms established during the engineering phase (Guarino et al., 2009).

An advanced supply chain tracking makes use of a sophisticated cross domain traceability and sensory mapping query. This is a very advanced industrial use case for this search across different conceptual domains. Starting with the macro geographic provenance, zooming in to the nation of Ethiopia, traveling through the physical agricultural bean entities associated with that origin and ending deep within the subjective human sensory evaluation profiles. The query returns highly quantitative, numeric scoring metrics on perceived acidity and physical body and qualitative, unstructured textual descriptors of flavor notes. It lets quality assurance professionals mathematically link precise geopolitical origins to very specific chemical sensory results across massive data sets.

Finally, a semantic unstructured text search query is the bridge between formalized mathematical logic and highly variable unstructured human vocabulary. This execution path employs advanced string manipulation and filtering functions to interpret the textual flavor notes attribute. This tool enables fuzzy, pattern-matching search of coffee lots with specific aromatic chemical compounds, like floral notes, irrespective of botanical species, processing method or geographic region. This is an important capability for retail stakeholders who wish to curate products based on consumer flavor preferences rather than technical agricultural specifications.

The widespread industrial deployment and use of formal ontologies is highly dependent on the availability of very user-friendly, computationally scalable software tools that successfully hide the extreme complexity of raw RDF and XML syntax (Coene, 2018; Chart.js documentation, 2026; Lavanya et al., 2023). The web-based exploration application was developed from scratch to directly

address this critical industry need. It is a very modern Single Page Application. It intentionally drops all heavy, backend JVM based parsing pipelines and uses a very decentralized, purely client-side processing architecture.

The application is built on React.js framework and is written in modern day JavaScript standards, HTML5 and CSS3. Responsive styling is supported via a modern user interface framework. React's component-driven, state-reactive paradigm is a natural fit for the complex, deeply nested graph data of parsed ontologies, which can be mapped directly and efficiently to discrete, reusable UI elements. The software project directory is organized into separate logical layers in an effort to maintain the codebase for a long period of time.

The contexts layer handles global application state and uses sophisticated hook patterns to offer very secure user authentication sessions and instant real-time bilingual localization without the highly inefficient component prop-drilling methodology. The component layer hosts the modular interface elements, defines the overall layout, handles the cryptographic authentication flows and contains the core exploratory functionalities. Finally, the critical utility layer comprises the core algorithmic parsing logic, most notably the proprietary parsing engine, which handles in a stand-alone fashion the complex ingestion, structural transformation and state integration of raw ontological data.

A new technical innovation central to this software architecture is the complete implementation of deep ontology parsing within the isolated context of the user's local web browser. Traditional semantic web applications inherently work by posting raw files to a remote, highly provisioned server with heavy Java-based endpoints (Dudáš et al., 2018; Katifori et al., 2007). This traditional client-server approach, while being computationally powerful for massive datasets, suffers from significant network latency, the need for continuous maintenance of expensive backend infrastructure and a fundamental compromise of data security in dealing with highly proprietary supply chain data.

The implemented exploration tool overcomes these deep limitations by deploying a custom parsing algorithm that uses the browser's native document object model parsing application programming interface strictly. The algorithmic flow is based on a very sequential pipeline. The user first uploads a raw ontology file, which is read directly into local system memory as an unformatted raw text string. Then the parsing engine is run, and a huge Abstract Syntax Tree (AST) is instantly created, which maps the whole XML hierarchy in memory.

The sequential semantic extraction algorithms traverse the nodes actively after the syntax tree generation. The algorithms specifically select out structural namespaces. The document is queried for defined class nodes and both object and datatype properties are recursively extracted, capturing exactly their respective domain and range mathematical constraints. In the final algorithmic phase, an enrichment function converts the raw, extracted XML nodes into a very optimized, deeply nested JSON data structure that is specifically designed for very fast traversal and reconciliation by React's internal state management hooks. This purely client-side approach significantly improves corporate data privacy by never transmitting proprietary industrial ontologies outside of the local computational machine, while at the same time increasing the responsiveness of the interface by offloading the entire computational workload to the client's local processing unit instead of remote servers (Bach et al., 2011; Dudáš et al., 2018).

3. Results

True native execution of semantic query languages is technically impossible in the browser environment by deliberate design without a heavy backend graph database. To get around this hard limit in a programmatic manner, the software provides a sophisticated query simulation module. It is an intermediary interpreter that takes user-selected granular parameters via the graphical interface, dynamically generates highly equivalent JavaScript array-filtering logic, and executes those mathematical filters quickly and directly against the JSON-transformed ontology state currently held in local memory. The module algorithmically synthesizes the raw, mathematically correct query syntax string for educational viewing and external export, providing users with the exact code they need should they later choose to deploy the ontology on a dedicated enterprise server.

The visualization layer integrates highly optimized SVG based charting libraries for analysis of aggregate, macro level data. Complex data preparation algorithms dynamically aggregate instance counts, compute hierarchical distributions and compute certain property values. The transformed data is fed to reactive graphical charts immediately. This seamless pipeline transforms highly abstract, deeply mathematical semantic web data into immediately actionable, visually digestible business intelligence dashboards, perfect for non-technical industry stakeholders and supply chain managers.

The overall success of the proposed semantic solution was systematically assessed along two very different computational dimensions: the mathematical structural integrity of the generated semantic model itself and the raw computational performance indicators of the React based web exploration architecture.

A final confirmation that the developed domain ontology is a very robust interconnected knowledge graph and not a simple flat taxonomy was given by systematic use of established mathematical evaluation metrics (Maynard et al., 2006).

The resulting semantic model based on the well-documented internal architecture consists of exactly 27 conceptual classes altogether, both foundational base classes and mathematically derived classes. Besides, the model defines an exact number of 21 different semantic properties, all of them strictly divided into 11 different

object properties and 10 very specific data type properties.

The most important metric of evaluation used is Relationship Richness. This metric represents a mathematical measure of total semantic relations diversity defined in the schema. The higher ratio decisively shows that the engineered ontology contains highly complex, domain-specific interactions, far beyond simple hierarchical parent-child inheritance structures (Burton-Jones et al., 2005; Maedche and Staab, 2002; Tartir et al., 2005). The calculation is based on the formal mathematical definition of the proportion of non-inheritance relationships to the sum of inheritance and non-inheritance relationships.

Considering a strict normalized tree structure, the number of inheritance relationships for the 27 defined classes is about 26. Mathematically, with the explicit use of the 11 object properties, the computation yields a ratio of 11/37 which yields a Relationship Richness indicator of exactly 0.297. Values close to 0.3 can be obtained, which indicates a very interconnected, very dense semantic network. Basic flat taxonomies mathematically tend towards an indicator of zero. This ontology tightly links distinct cross-domain nodes unequivocally confirming its high utility for complex supply chain traceability and sophisticated algorithmic recommendation systems (Tartir et al., 2005).

In ontology quality benchmarking (Tartir et al., 2005; Burton-Jones et al., 2005), Relationship Richness (RR) reflects the proportion of non-taxonomic relationships relative to the total number of relationships $\left(RR = |P_{non-isa}|/(|P_{isa}| + |P_{non-isa}|)\right)$. Purely taxonomic hierarchies (like simple thesauri or taxonomies) exhibit an RR near 0.0, as they rely almost exclusively on subClassOf inheritance. In domain-specific ontologies, RR values typically range between 0.10 and 0.35 depending on graph density. Our achieved RR of 0.297 mathematically confirms that nearly 30% of all defined relationships in the schema express complex, cross-domain functional associations (e.g., agronomic, chemical, and sensory linkages) rather than simple taxonomic subclassing, placing it at the upper end of domain-specific semantic networks.

Also Attribute Richness metric is evaluated in the study. This metric computes the average size of literal, descriptive data assigned per class (Burton-Jones et al., 2005; Maedche and Staab, 2002; Tartir et al., 2005). Dividing the 10 defined datatype properties by the total 27 classes an Attribute Richness indicator of 0.370 is obtained. This metric specifically reveals a very moderate, optimally balanced distribution of datatype properties that captures critical physical properties, accurate environmental temperatures and highly subjective sensory scores adequately and efficiently, to fully enable deep, quantitative machine queries without unnecessarily overloading the data structure.

Strict performance stress tests of the raw computational efficiency and the overall stability of the web-based exploration software conclusively proved the technical feasibility of the purely client-side document parsing architecture.

The application was packed heavily and optimized by advanced compiling techniques, achieving a final payload size of about 2.5 megabytes. This footprint was then compressed even further to a very efficient 800 kilobytes using gzip encoding. This ultra-lightweight footprint, which employs only sophisticated code splitting and very dynamic route-based lazy loading techniques and nothing else, provides for incredibly fast deployment and execution even over the very constrained, low-bandwidth networks common in rural agricultural environments.

Empirical runtime metrics show remarkable, uninterrupted fluidity in all operations. The app conclusively achieves a full initial load time strictly less than 3 seconds from cold start. The proprietary parsing engine takes about 500 milliseconds to process and build the AST for a medium-sized ontology file. Complex query simulations and search index operations also produce full responses in less than 100 milliseconds and the translation and rendering of complex SVG graphical data use less than 200 milliseconds of main thread execution time.

To make a definitive assessment of long-term scalability, the system architecture was aggressively stress-tested against massive, synthetically generated ontologies containing well over 1,000 distinct, heavily interlinked classes and properties. The huge memory allocation that was needed for the extensive syntax tree was handled efficiently and safely by the browser's native parsing API, without triggering the garbage collection stutters. The React virtual DOM also reconciled massive localized state changes at a very high efficiency rate, without blocking the main browser thread or visibly degrading the search input responsiveness, at the same time. These final empirical results provide strong supporting evidence for the core assumption that contemporary client browser engines have enough computational overhead to independently process, traverse and visualize super-strong semantic models without reliance at all on traditional backend server infrastructure. Client-Side Performance Benchmarks under Standardized Environment can be seen on Table 5.

Table 5. Client-side performance benchmarks under standardized environment*

Benchmark Parameter	Operational Value / Metric	Execution Context & Hardware Conditions
Base Ontology Size	27 classes, 21 properties, ~350 RDF triples (18 KB OWL/XML)	Baseline test file
Synthetic Stress Test Dataset	1,024 classes, 85 properties, ~12,500 RDF triples (1.4 MB)	Scalability stress test
Initial Bundle Footprint	800 KB (Gzip compressed) / 2.5 MB uncompressed	Optimized React production build
AST Generation Time	~500 ms (Baseline) / ~1,850 ms (12k triples)	Native Browser DOM Parser
Query Sim / Filtering	< 100 ms	Client-side JS array manipulation
SVG Render Latency	< 200 ms	React Virtual DOM reconciliation

*Test Environment Hardware Setup for Reproducibility: Intel Core i7-12700H @ 2.30GHz, 16 GB DDR4 RAM, Google Chrome v125.0 (64-bit) running on Windows 11 Enterprise. Network latency was throttled using Chrome DevTools to simulate low-bandwidth rural connections.

4. Discussion

The deep semantic ontology is tightly integrated with the high-performance, React-based web exploration architecture and has several immediate and profound implications for the wider agricultural technology space. By changing the huge vocabulary of the global coffee supply chain dramatically from highly disparate, standalone relational databases to a unified, mathematically rigorous Semantic Web model, industry stakeholders will be able to achieve truly unprecedented end-to-end traceability (Ramos et al., 2023; Mauladi et al., 2022).

For instance, the strategic combination of such semantic ontology with physical IoT sensors, deployed directly on the agricultural farm, along with immutable blockchain logistics platforms, guarantees that highly critical data is not only stored in a database, but is inherently, contextually understood by fully automated reasoning systems (Kamble et al., 2020; Lin et al., 2020; Ramos et al., 2019). The strict mathematical constraints that are built directly into the ontology allow fully automated, highly trusted certification processes, massively reducing manual administrative overhead, removing human error and completely minimizing the risk of fraudulent environmental or quality labeling within international trade.

From a strict software engineering perspective, the shown performance of the client-side parsing mechanism suggests a fundamental, larger change in the way in which Semantic Web tools should be designed and deployed in the future. Traditional architectural dependencies on heavy, backend Java-based parsing endpoints inherently create massive bottlenecks in user experience development, drastically increase corporate hosting costs, and severely slow down the widespread use of ontology exploration outside of isolated academic research circles (Horridge, 2007; Katifori et al., 2007). The very novel architecture presented in this comprehensive study clearly demonstrates that highly complex analytical tools, no matter how heavy, can be fully distributed through simple static file hosting, thus finally democratizing access to highly complex formalized knowledge structures for all levels of users.

The present version of the software architecture, however, has some technical limitations which clearly point to the direction of future research and development. The current implementation is very performant when analyzing files locally, but it relies on an emulated query engine through JavaScript array manipulations, which mathematically cannot substitute a true graph database with highly distributed, federated querying capabilities over massive, petabyte-scale datasets.

The proposed Coffee Culture and Production Processes Ontology has also to be maintained and developed over time beyond its immediate use, which is a key aspect to keep its academic and industrial relevance. The agricultural domain is dynamic with new botanical cultivars, processing technologies and sensory evaluation standards being introduced continuously. Therefore, the semantic model cannot be static. A formalized maintenance workflow is expected to keep on the continuous vitality. First, the ontology will be hosted in a public, version-controlled repository (e.g. GitHub) so that the community of agronomy researchers and software developers are able to submit pull requests for new concepts or property alignments. Second, future incarnations will follow established semantic versioning practices (such as OWL's owl:versionIRI properties) to ensure backward compatibility with legacy data and traceability systems that already depend on earlier drafts. Finally, we will link our micro-ontology to developing foundational macro-ontologies such as FoodOn and AGROVOC using explicit mapping properties (e.g., owl:equivalentClass or skos:exactMatch) to reduce the risk of conceptual drift and the burden of manual curation. This is a community-driven, standards-aligned model for maintenance, turning the ontology from a static snapshot into a living, community-maintained semantic standard for the global coffee industry.

Furthermore, in future extensions of this research, incorporating state-of-the-art Deep Learning (DL)

methodologies presents a promising direction. Deep learning models (e.g., Graph Neural Networks for knowledge graph embeddings or transformer-based architectures for automated entity/relation extraction from unstructured agricultural texts) could be explored to semi-automatically populate and evolve the coffee ontology. Combining symbolic reasoning from our OWL ontology with subsymbolic representations from deep learning paradigms could further enhance predictive modeling for sensory evaluation and crop yield optimization.

Future versions of the architecture will focus on the implementation of extremely advanced, hybrid architectures, where the client application will be able to connect seamlessly to an external, real-time query endpoint, while keeping the highly efficient local parsing power for ad-hoc, offline file exploration. Moreover, the future integration of advanced Large Language Models to successfully enable pure natural language querying over the complex ontology via localized Retrieval-Augmented Generation paradigms will serve to further lower the technical barrier to entry for non-technical agricultural users, completely revolutionizing how human operators interact with mathematically formalized semantic data (Lippolis et al., 2025; Tiwari et al., 2025; Sharma et al., 2024).

5. Conclusion

The rapid and broad acceptance of highly standardized, machine-readable knowledge representations is an absolute necessity for the seamless, error-free operation of modern global agricultural supply chains. This extensive research resulted in a highly relational, very complex domain ontology explicitly developed for the complexity of coffee culture and precision production. This semantic model was rigorously validated by a quantitative Relationship Richness metric that indicated very high semantic density and advanced inferential capabilities in an automated way. At the same time, the successful architectural development and deployment of the web-based exploration tool establish a highly performant, fully serverless architectural pattern for web-based semantic visualization of data. The framework presented here successfully closes the historical gap between highly formal, mathematical ontology engineering and accessible, user-centric commercial software design, by decisively demonstrating sub-second document parse times and truly instantaneous query responses using a native, client-side data parsing pipeline. To summarize, this dual computational approach provides the absolutely essential foundational infrastructure needed to accelerate large-scale data harmonization, automated quality assurance, and unbreakable traceability in the highly complex global coffee industry.

Author Contributions

The percentages of the authors' contributions are presented below. All authors reviewed and approved the final version of the manuscript.

	Ö.D.	A.D.
C	50	50
D	40	60
S	0	100
DCP	80	20
DAI	70	30
L	70	30
W	50	50
CR	20	80
SR	20	80
PM	10	90
FA	0	0

C= concept, D= design, S= supervision, DCP= data collection and/or processing, DAI= data analysis and/or interpretation, L= literature search, W= writing, CR= critical review, SR= submission and revision, PM= project management, FA= funding acquisition.

Conflict of Interest

The authors declared that there is no conflict of interest.

Ethical Consideration

Ethics committee approval was not required for this study because of there was no study on animals or humans.

References

Angelica, M., & Ferdinand, F. N. (2017). Expert system based on an ontology method to analyze types of Arabica coffee beans. *International Journal of Interactive Mobile Technologies*, *11*(4), 1–10. https://doi.org/10.3991/ijim.v11i4.6859

Bach, B., Pietriga, E., Liccardi, I., & Legostaev, G. (2011). OntoTrix: A hybrid visualization for populated ontologies. In *Proceedings of the 20th International Conference Companion on World Wide Web* (pp. 177–180). Association for Computing Machinery. https://doi.org/10.1145/1963192.1963282

Burton-Jones, A., Storey, V. C., Sugumaran, V., & Ahluwalia, P. (2005). A semiotic metrics suite for assessing the quality of ontologies. *Data & Knowledge Engineering*, *55*(1), 84–102. https://doi.org/10.1016/j.datak.2004.11.010

Chart.js. (n.d.). *Chart.js documentation*. Retrieved June 16, 2026, from https://www.chartjs.org/docs/latest

Coene, J. P. (2018). Sigmajs: An R htmlwidget interface to the sigma.js visualization library. *Journal of Open Source Software*, *3*(28), Article 814. https://doi.org/10.21105/joss.00814

Dooley, D. M., Griffiths, E. J., Gosal, G. S., Buttigieg, P. L., Bastien, R., Jalali, N. N., Hoehndorf, R., Schriml, L. M., De Coronado, A. R., & Hsiao, W. W. L. (2018). FoodOn: A harmonized food ontology to increase global food traceability, quality control and data integration. *npj Science of Food*, *2*, Article 23. https://doi.org/10.1038/s41538-018-0032-6

Dudáš M, Lohmann S, Svátek V, Pavlov D. Ontology visualization methods and tools: a survey of the state of the art. *The*

Knowledge Engineering Review. 2018;33:e10. doi:10.1017/S0269888918000073

Goldstein, A., Fink, L., & Ravid, G. (2021). A framework for evaluating agricultural ontologies. *Sustainability*, *13*(11), Article 6387. https://doi.org/10.3390/su13116387

Gruber, T. R. (1993). A translation approach to portable ontology specifications. *Knowledge Acquisition*, *5*(2), 199–220. https://doi.org/10.1006/knac.1993.1008

Guarino, N., Oberle, D., & Staab, S. (2009). What is an ontology? In S. Staab & R. Studer (Eds.), *Handbook on ontologies* (2nd ed., pp. 1–17). Springer. https://doi.org/10.1007/978-3-540-92673-3_0

Horridge, M. (2007). *Pizza ontology 1.5 tutorial*. Manchester University.

Illy, A., & Viani, R. (Eds.). (2005). *Espresso coffee: The science of quality* (2nd ed.). Academic Press.

Joo, S., Koide, S., Takeda, H., Horyu, D., Takezaki, A., & Yoshida, T. (2016). Designing of ontology for domain vocabulary on Agriculture Activity Ontology (AAO) and a lesson learned. In Y.-F. Li, M. M. S. Wong, G. Qi, & Y. Zou (Eds.), *Semantic technology. JIST 2016* (Lecture Notes in Computer Science, Vol. 10055, pp. 31–43). Springer, Cham. https://doi.org/10.1007/978-3-319-50112-3_3

Kamble, S. S., Gunasekaran, A., & Sharma, R. (2020). Modeling the blockchain enabled traceability in agriculture supply chain. *International Journal of Information Management*, *52*, Article 101967. https://doi.org/10.1016/j.ijinfomgt.2019.05.023

Katifori, A., Halatsis, C., Lepouras, G., Vassilakis, C., & Giannopoulou, E. (2007). Ontology visualization methods—A survey. *ACM Computing Surveys*, *39*(4), Article 10. https://doi.org/10.1145/1287620.1287621

Keet, M. (2019). *An introduction to ontology engineering*. College Publications.

Lavanya, A., Sindhuja, S., Gaurav, L., & Ali, W. (2023). A comprehensive review of data visualization tools: Features, strengths, and weaknesses. *International Journal of Computer Engineering and Research Trends*, *10*(1), 10–20.

Leung, C., & Salga, A. (2010). Enabling WebGL. In *Proceedings of the 19th International Conference Companion on World Wide Web* (pp. 10–14). Association for Computing Machinery. https://doi.org/10.1145/1772690.1772933

Liang, A., Sini, M., Lu, W., Keizer, J., & Katz, S. (2006). *The mapping schema from Chinese agricultural thesaurus to AGROVOC*. Food and Agriculture Organization of the United Nations.

Lin, Y. P., Petway, J. R., Anthony, J., Mukhtar, H., Liao, S. W., Chou, C. F., & Ho, Y. F. (2020). Blockchain: The evolutionary next step for ICT e-agriculture. *Environments*, *4*(3), Article 50. https://doi.org/10.3390/environments4030050

Lippolis, A. S., Siragusa, G., Buscaldi, D., & Rossi, R. (2025). *Ontology generation using large language models* (arXiv:2503.05388). arXiv. https://doi.org/10.48550/arXiv.2503.05388

Lohmann, S., Negru, S., Haag, F., & Ertl, T. (2014). Visualizing ontologies with VOWL. *Semantic Web*, *5*(5), 399–419. https://doi.org/10.3233/SW-140134

Maedche, A., & Staab, S. (2002). Measuring similarity between ontologies. In A. Gómez-Pérez & V. R. Benjamins (Eds.), *Knowledge engineering and knowledge management: Ontologies and the Semantic Web. EKAW 2002* (Lecture Notes in Computer Science, Vol. 2473, pp. 251–263). Springer, Berlin, Heidelberg. https://doi.org/10.1007/3-540-45810-7_24

Mauladi, M. A. R., Jangkung Handoyo Mulyo, & Dwidjono Hadi Darwanto. (2022). Coffee Supply Chain Management: A Case Study In Ciamis, West Java, Indonesia. *HABITAT*, *33*(3), 201–211. https://doi.org/10.21776/ub.habitat.2022.033.3.20

Maynard, D., Peters, W., & Li, Y. (2006). Metrics for evaluation of ontology-based information extraction. In *Proceedings of the WWW 2006 Workshop on Evaluation of Ontologies for the Web (EON 2006)* (pp. 1–8). CEUR Workshop Proceedings.

Noy, N. F., & McGuinness, D. L. (2001). *Ontology development 101: A guide to creating your first ontology* (Stanford Knowledge Systems Laboratory Technical Report KSL-01-05). Stanford University.

Nuraisyah, A., Wulandari, E., Indrawan, D. *et al.* The roles of stakeholders in supply chain sustainability challenges: the case of coffee chain in West Java Province, Indonesia. *Discov Sustain* **6**, 247 (2025). https://doi.org/10.1007/s43621-025-01004-3

Ramos, E., Chavez, M., Patrucco, A. S., & Gonzales, A. (2019). Organic coffee supply chain source process integration: A Peruvian case. *International Journal of Supply Chain Management*, *8*(6), 133–145.

Ramos, E., Patrucco, A. S., & Chavez, M. (2023). Dynamic capabilities in the "new normal": A study of organizational flexibility, integration and agility in the Peruvian coffee supply chain. *Supply Chain Management: An International Journal*, *28*(1), 55–73. https://doi.org/10.1108/SCM-12-2021-0553

Sharma, K., Kumar, P., & Li, Y. (2024). *OG-RAG: Ontology-grounded retrieval-augmented generation for large language models* (arXiv:2412.15235). arXiv. https://doi.org/10.48550/arXiv.2412.15235

Smaili, N., & Kabbaj, A. (2025). Enabling semantic interoperability for smart farming. *Agronomy Research*, *23*(1), 479–492. https://doi.org/10.15159/AR.25.021

Tartir, S., Arpinar, I. B., Moore, M., Sheth, A. P., & Aleman-Meza, B. (2005). OntoQA: Metric-based ontology quality analysis. In *Proceedings of the IEEE Workshop on Knowledge Acquisition from Distributed, Autonomous, Semantically Heterogeneous Data (KADASH 2005)* (pp. 1–8). IEEE.

Tiwari, Y., Lone, O. A., & Pal, M. (2025). *OntoRAG: Enhancing question-answering through automated ontology derivation from unstructured knowledge bases* (arXiv:2506.00664). arXiv. https://doi.org/10.48550/arXiv.2506.00664

W3C OWL Working Group. (2012). *OWL 2 Web Ontology Language document overview* (2nd ed., W3C Recommendation). W3C. https://www.w3.org/TR/owl2-overview/

W3C SPARQL Working Group. (2013). *SPARQL 1.1 query language* (W3C Recommendation). W3C. https://www.w3.org/TR/sparql11-query/