© SEP 2026 | IRE Journals | Volume 10 Issue 3 | ISSN: 2456-8880
DOI: https://doi.org/10.64388/IREV10I3-1723181

Data-Driven Identification of Industrial Localization Opportunities in Saudi Arabia: A Big Data Framework for Vision 2030

SYED MOHAMMAD ALI

Abstract-** Saudi Arabia is undertaking a large-scale economic and industrial transformation through Vision 2030, with industrial localization, local-content development, supply-chain resilience and non-oil growth forming important components of the national agenda. Yet localization decisions are often complicated by fragmented information across customs records, industrial production, procurement, company registries, supplier capabilities, technology requirements and future investment pipelines. This paper develops a data-driven framework for identifying and prioritizing industrial localization opportunities in Saudi Arabia through the integration of big data analytics, trade intelligence, demand modelling, supplier-ecosystem assessment, value-chain analysis and multi-criteria decision support. The study adopts a structured review and conceptual-framework methodology and synthesizes recent literature with official Saudi policy and trade evidence. Six analytical dimensions are proposed: import dependency, market attractiveness, supplier capability, value-chain localization potential, strategic importance and economic impact. These dimensions are combined through a transparent Localization Opportunity Index (LOI), which can be applied at product, industry or value-chain level. The paper further outlines a practical data architecture incorporating data engineering, SQL/Python analytics, business-intelligence dashboards, forecasting, machine learning and network analysis. The framework demonstrates why import value alone is an insufficient basis for localization decisions and shows how product-level, supplier-level and value-chain-level analytics can improve industrial prioritization. For Saudi Arabia, the approach can support more evidence-based investment attraction, local-content development, industrial resilience, SME participation and technology transfer under Vision 2030. The paper concludes with implementation, governance and research recommendations for building a continuously updated localization decision-support capability.

***Keywords:** Saudi Vision 2030; industrial localization; big data analytics; local content; supply chain; import dependency; supplier ecosystem; value-chain analysis; industrial policy; data-driven decision-making.*

I. INTRODUCTION

Saudi Arabia's Vision 2030 has accelerated the Kingdom's transition toward a more diversified, investment-driven and knowledge-intensive economy. Industrial development is central to this transition because competitive manufacturing can generate domestic value addition, technological capability, skilled employment, export diversification and stronger upstream and downstream supplier networks. The Vision explicitly highlights the localization of promising manufacturing industries and the development of a more competitive non-oil economy (Saudi Vision 2030, 2026). The National Industrial Strategy similarly emphasizes expansion of local production, resilient supply chains and higher-value industrial activity.

Localization is more complex than simply replacing imports. An economy benefits from international trade, and many imported products may remain more efficient to source globally. Strategic localization therefore requires a disciplined method for distinguishing between products that should continue to be imported and products for which domestic production can generate sustainable economic, technological or resilience benefits. Such decisions must consider the size and stability of demand, existing domestic capacity, availability of raw materials, supplier readiness, technology requirements, investment intensity, workforce skills, future project pipelines and export potential.

The scale of Saudi Arabia's merchandise trade demonstrates the magnitude of the analytical challenge. According to the General Authority for Statistics (GASTAT), total merchandise exports reached SAR 90.4 billion in May 2025, merchandise imports reached SAR 80.9 billion, and non-oil exports including re-exports reached SAR 31.1 billion. These

© SEP 2026 | IRE Journals | Volume 10 Issue 3 | ISSN: 2456-8880
DOI: https://doi.org/10.64388/IREV10I3-1723181

values are not evidence that all imports should be localized; rather, they show that the Kingdom has a large and diverse trade base within which strategically attractive product-level opportunities can be identified (GASTAT, 2025a).

Conventional industrial screening commonly relies on trade tables, expert workshops and spreadsheet-based rankings. These methods remain useful, but they become increasingly difficult to scale when analysts must compare thousands of product codes, hundreds of industries, multiple years of trade transactions, supplier databases, corporate profiles and project-level demand. Big data analytics can address this challenge by integrating structured and semi-structured datasets, identifying patterns, forecasting demand, mapping supplier networks and translating complex evidence into transparent decision scores.

The relevance of this approach is reinforced by the author's professional domain. His experience includes integrating complex trade, product, demand, company and supplier datasets; screening industrial sectors for localization potential; conducting value-chain mapping; developing demand and impact models; and assessing supplier readiness. This professional context motivates the research question while the paper itself remains a conceptual and evidence-based scholarly contribution rather than a case report.

This paper therefore asks: How can big data analytics be systematically applied to identify and prioritize industrial localization opportunities in Saudi Arabia in support of Vision 2030? To answer this question, the study develops a multidimensional framework that combines trade dependence, market attractiveness, supplier capability, value-chain feasibility, strategic importance and economic impact. It also proposes an implementation architecture that links data engineering, analytical modelling and business-intelligence visualization with institutional governance and expert judgment.

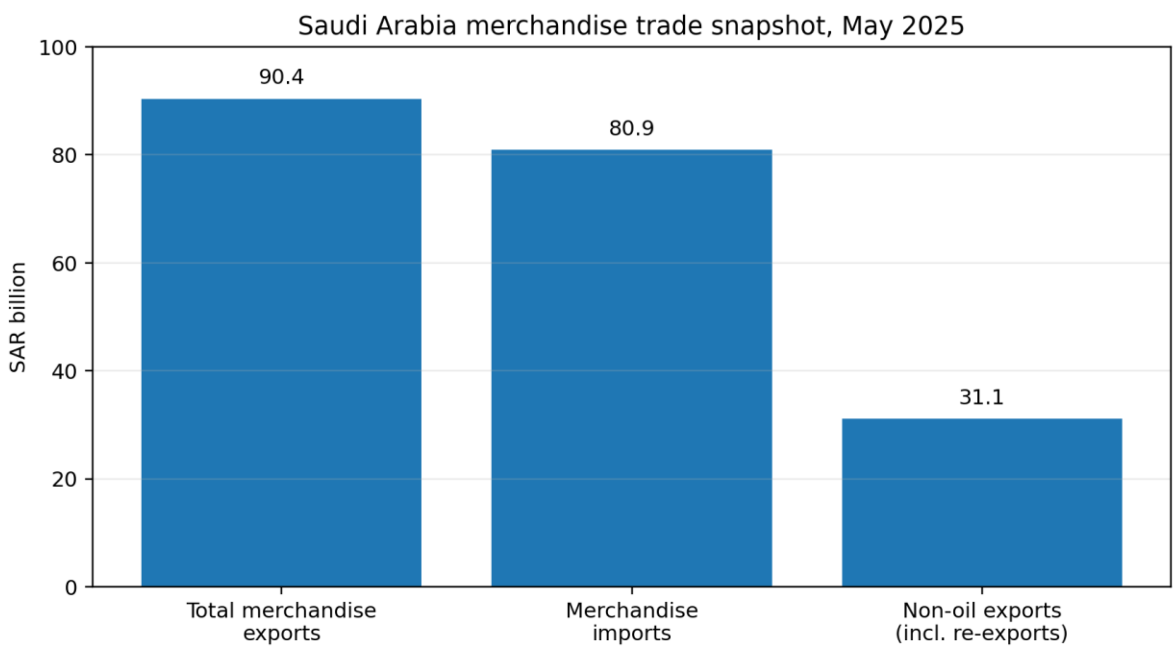

Figure 1. Saudi Arabia merchandise trade snapshot, May 2025. Source: GASTAT (2025a).

1.1 Research Aim

The aim of this study is to develop a big-data-driven analytical framework for systematically identifying, evaluating and prioritizing industrial localization opportunities in Saudi Arabia under Vision 2030.

1.2 Research Objectives

- Examine the role of industrial localization and local-content development within Saudi Arabia’s economic transformation.

- Assess how big data analytics can improve the quality and speed of localization opportunity identification.
- Identify the core datasets and analytical dimensions required for national and sector-level localization screening.
- Develop a transparent Localization Opportunity Index (LOI) for comparative product and industry prioritization.
- Integrate supplier-ecosystem and value-chain analysis into localization feasibility assessment.
- Evaluate implementation challenges related to data quality, governance, skills, institutional coordination and model explainability.
- Provide practical recommendations for policymakers, industrial-development organizations, investors and large Saudi buyers.

1.3 Research Contribution

The study contributes to the industrial-policy and digital-transformation literature by treating localization as a data-integration and decision-intelligence problem. It moves beyond conventional import-substitution analysis in three important ways. First, it separates import exposure from localization attractiveness: a product can have high imports but low localization feasibility. Second, it incorporates supply-side feasibility and supplier readiness rather than focusing only on market size. Third, it treats localization as a dynamic process that should be continuously monitored as demand, technology, prices, suppliers and national priorities change.

II. SAUDI INDUSTRIAL LOCALIZATION IN THE CONTEXT OF VISION 2030

Saudi Vision 2030 seeks to create a thriving economy with stronger private-sector participation, investment, innovation and diversification. The national agenda includes localization of promising manufacturing industries, development of mining and industrial value chains, digital-economy expansion and increased competitiveness of Saudi businesses. The National Industrial Development and Logistics Program (NIDLP) links industrial development with logistics, mining and energy, recognizing that competitive industry depends on integrated infrastructure, investment and supply networks.

Industrial localization contributes to economic diversification through several mechanisms. Domestic production can increase local value added by shifting selected activities from importation toward processing, manufacturing, assembly, engineering, testing, maintenance and product development. It can create jobs and deepen technical capabilities, particularly when localization programs are accompanied by supplier development, training and technology partnerships. It can also reduce exposure to highly concentrated foreign sources for strategically important products.

However, the strongest localization strategy is selective rather than protectionist. Products with limited domestic demand, very high technological barriers or weak access to essential inputs may not be economically suitable for domestic production. Conversely, products with large and growing demand, strong adjacent capabilities and significant strategic importance can become attractive even if initial production costs are higher than imported alternatives. The analytical task is therefore to identify where the balance of economic and strategic factors favors localization.

Saudi trade data also reveal an important dual trend: imports remain substantial while non-oil exports are expanding. GASTAT reported that non-oil exports including re-exports increased by 22.1% year-on-year in June 2025, while Q2 2025 non-oil exports rose by 17.8%. Machinery, electrical equipment and parts represented 28.9% of total imports in Q2 2025 and 21.7% of non-oil exports, while chemical products accounted for 23.0% of non-oil exports (GASTAT, 2025b). These patterns show both the importance of import-dependent industrial categories and the possibility of developing competitive non-oil export capabilities.

A modern localization agenda therefore needs to answer not only “what do we import?” but also “which products are strategically material, which domestic firms can participate, which value-chain stages can be localized, how will demand evolve, what economic impact can be generated, and what type of intervention is appropriate?” Big data analytics provides a mechanism for answering these questions consistently across thousands of potential opportunities.

III. LITERATURE REVIEW

3.1 Big Data Analytics and Industrial Decision-Making

Big data analytics refers to the use of advanced computational and statistical techniques to derive insight from datasets characterized by high volume, variety, velocity and complexity. In industrial settings, relevant data can include customs transactions, product classifications, supplier catalogues, production records, procurement histories, company financials, investment announcements, patent records, logistics movements and technical specifications. The analytical value of these data emerges when they are integrated rather than interpreted in isolation.

Four broad forms of analytics are relevant. Descriptive analytics explains what has happened, such as the value and growth of imports by product code. Diagnostic analytics investigates why patterns exist, including supplier concentration or recurring domestic capability gaps. Predictive analytics estimates future conditions using historical trends and explanatory variables. Prescriptive analytics supports the selection of interventions or investments by combining forecasts with explicit objectives and constraints.

Recent research in Saudi and regional contexts has increasingly focused on the relationship between digital capabilities, supply-chain performance and sustainability. Studies of Saudi organizations indicate that technological readiness, data quality, skills and organizational capabilities shape the value generated from analytics. This implies that a national localization platform cannot be built solely as a software project; it requires institutional ownership, expert interpretation and disciplined data governance. Big data is particularly relevant to localization because the problem contains thousands of possible product-level decisions. Even within a single sector, products differ substantially in market scale, complexity, raw-material dependency, regulation and supplier maturity. Automated data pipelines can first screen and rank a very large opportunity universe, after which sector experts can conduct deeper technical and commercial validation.

3.2 Digital Transformation, Industry 4.0 and Localized Manufacturing

Industry 4.0 technologies—including industrial Internet of Things systems, artificial intelligence, cloud platforms, digital twins, advanced robotics, additive manufacturing and real-time analytics—are changing the economics of manufacturing. Flexible and digitally controlled production can reduce setup times, improve quality, support shorter product runs and enable greater customization. For some product categories, this can lower the minimum efficient scale needed for localization.

Aljuaid (2024) examined the integration of Industry 4.0 in sustainable localized manufacturing, emphasizing the role of digital technologies and big data in improving localized production models. Such findings are relevant to Saudi Arabia because localization feasibility is not static: an opportunity considered uneconomic under conventional production may become viable when advanced manufacturing, automation or digital design reduces cost and increases flexibility.

Digital transformation also strengthens supply-chain visibility. When procurement, inventory, supplier performance and logistics information are captured digitally, firms can identify dependency risks and develop alternative sourcing strategies more rapidly. Therefore, industrial digitalization should be considered both an enabler of manufacturing competitiveness and a source of better localization intelligence.

3.3 Supply-Chain Resilience and Strategic Localization

Supply-chain resilience is the ability of a supply network to anticipate, absorb, recover from and adapt to disruption. Geographic concentration, single-source dependency, long lead times, poor visibility and limited alternative capacity can increase vulnerability. Localization can reduce selected vulnerabilities by developing domestic or regional alternatives, particularly for critical inputs. Yet domestic production is not automatically resilient if the local producer remains dependent on imported raw materials, proprietary technology or a single overseas component.

This observation makes value-chain analysis essential. A localization project should be evaluated not only by the location of final assembly but by the share of economic and technical value that can be performed locally. A product assembled in Saudi Arabia from almost entirely imported components may create some jobs and logistics benefits but could still have limited resilience impact. By contrast, an ecosystem that develops local materials, components, engineering services and maintenance capability generates deeper localization.

Recent supply-chain research in the MENA region gives increasing attention to resilience, sustainability and digital transformation. These themes support the development of a Saudi localization model that explicitly combines economic opportunity with strategic vulnerability and digital capability.

3.4 Research Gap

Existing research provides substantial insight into industrial digitalization, supply-chain resilience and economic diversification, but fewer studies provide an integrated, product-level analytical framework designed specifically to prioritize industrial localization opportunities in Saudi Arabia. The gap is not the absence of trade data or industrial policy; it is the absence of a transparent architecture that links trade dependency, demand, supplier readiness, value-chain depth, strategic importance and economic impact within one repeatable decision process. This paper addresses that gap.

IV. RESEARCH METHODOLOGY

4.1 Research Design

The paper adopts a structured review and conceptual-framework methodology. It synthesizes recent peer-reviewed literature on big data analytics, Industry 4.0, supply-chain resilience and industrial localization with official Saudi Vision 2030 and GASTAT evidence. The purpose is not to claim results from a confidential national customs or supplier database; instead, the study develops a replicable analytical architecture that can be implemented when the required datasets are available.

The methodology follows five stages: (1) literature and policy synthesis; (2) identification of relevant industrial datasets; (3) definition of localization assessment dimensions; (4) construction of the Localization Opportunity Index; and (5) development of implementation, governance and policy recommendations. This design is appropriate for a framework article because it makes assumptions explicit and creates a basis for subsequent empirical validation.

4.2 Data Architecture and Candidate Data Sources

A national or enterprise-level localization platform requires data from multiple sources because no single dataset can capture both market attractiveness and production feasibility. Table 1 summarizes the major categories of data that can support localization screening.

Table 1. Data categories for industrial localization analytics.

Data category	Illustrative variables	Localization purpose
International trade	Import value/quantity, HS code, origin, unit value, export value	Measure dependency, market size and source concentration
Domestic production	Output, capacity, industrial licenses, utilization	Identify existing supply and capacity gaps
Procurement	Buyer, product, contract value, frequency, specifications	Validate demand and buyer concentration
Company data	Sector, revenue, employment, location, ownership	Map buyers, producers and adjacent capabilities
Supplier data	Products, machinery, certifications, capacity, quality systems	Assess local supplier readiness

© SEP 2026 | IRE Journals | Volume 10 Issue 3 | ISSN: 2456-8880
DOI: https://doi.org/10.64388/IREV10I3-1723181

Data category	Illustrative variables	Localization purpose
Investment/project pipeline	Planned projects, capex, commissioning dates	Forecast future demand
Workforce/skills	Occupation, qualification, wage, availability	Evaluate human-capital feasibility
Technology/IP	Patents, process requirements, equipment, licenses	Estimate technology complexity and transfer needs
Logistics/geospatial	Ports, industrial cities, road/rail, customer clusters	Assess location and distribution feasibility
Macroeconomic/input-output	GDP, sector linkages, multipliers	Estimate direct and indirect economic impact

4.3 Data Preparation and Quality Control

Data preparation is a critical step because analytical models can produce misleading rankings when source data are inconsistent. Product-code standardization is required where trade records use Harmonized System (HS) codes but procurement systems use internal material codes or industry classifications. Crosswalk tables can connect different taxonomies and permit aggregation at multiple levels.

Entity resolution is equally important. A single company may appear under different legal names, trade names or abbreviations. Matching algorithms can use company registration numbers, addresses, domain names, phone numbers and similarity methods to produce unified entity records. Supplier and buyer profiles can then be connected to products and industries.

Outlier detection should distinguish true market events from reporting errors. A sharp one-year increase in a product's import value may reflect a major capital project rather than recurring annual demand. Time-series normalization should therefore identify one-time purchases, classification changes, inflation effects and unit-value anomalies. Missing supplier data should also be tracked because missing information can itself signal low data maturity or the need for direct validation.

4.4 Analytical Dimensions and Localization Opportunity Index

The proposed framework uses six primary analytical dimensions. Each dimension is converted to a normalized score between 0 and 100. Weightings can be adjusted by sector because, for example, a health-security product may require greater strategic weighting than a consumer product.

Table 2. Core dimensions of the Localization Opportunity Index.

Dimension	Example measures	Decision meaning
Import dependency (ID)	Import share of apparent consumption; import growth; source concentration	Exposure to foreign supply
Market attractiveness (MA)	Market size; demand growth; future projects; export potential	Commercial scale and growth
Supplier capability (SC)	Existing manufacturers; certifications; capacity; adjacent capabilities	Feasibility of local production
Value-chain potential (VC)	Local inputs; localizable stages; service ecosystem; technology access	Depth of localization

Dimension	Example measures	Decision meaning
Strategic importance (SI)	Criticality to health, food, energy, infrastructure or industry	National resilience value
Economic impact (EI)	Value added; jobs; GDP; SME participation; trade-balance impact	National development benefits

A generalized Localization Opportunity Index can be represented as: LOI = w1(ID) + w2(MA) + w3(SC) + w4(VC) + w5(SI) + w6(EI), where the w terms represent weights determined by policy or sector priorities. The principal advantage of the index is transparency: decision-makers can see why an opportunity ranks highly and can perform sensitivity analysis when weights change.

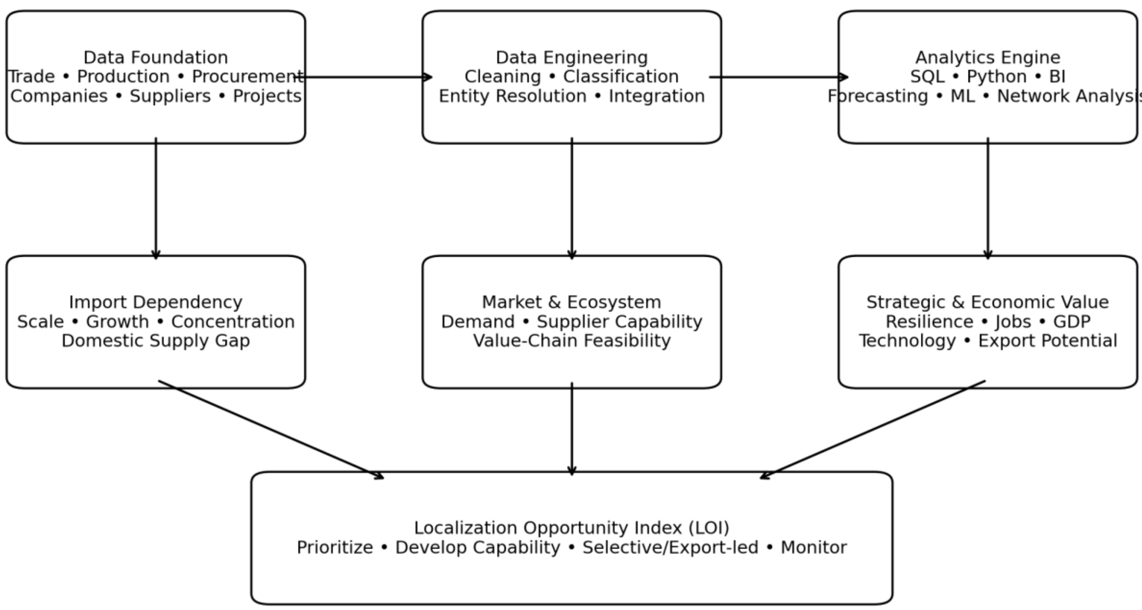

Figure 2. Proposed big-data framework for industrial localization opportunity identification. Source: Author-developed framework.

V. THE PROPOSED BIG DATA LOCALIZATION FRAMEWORK

5.1 Layer 1: Import-Dependency Analysis

Import-dependency analysis is the initial screening layer. Relevant measures include annual import value, multi-year import growth, import quantities, unit values, number of source countries, source-country concentration, number of major buyers and domestic production relative to apparent consumption. A high dependency score identifies categories where Saudi demand is substantially supplied from abroad.

A simple import-dependency ratio can be defined as Imports divided by apparent domestic consumption, where apparent consumption equals domestic production plus imports minus exports. The ratio is useful but should not be interpreted alone. A niche product may have a ratio close to 100% but a market too small to support competitive domestic production. Conversely, a product with moderate import dependency but rapidly expanding domestic demand may become a stronger opportunity.

Source concentration is another important variable. A product imported from many countries has different resilience characteristics from a product primarily dependent on one or two countries. Measures such as the Herfindahl-Hirschman Index can be adapted to quantify concentration among import origins or major suppliers.

5.2 Layer 2: Market-Attractiveness and Demand Forecasting

Market attractiveness determines whether sufficient and durable demand exists to justify localized investment. Historical customs data should be combined with forward-looking variables including population growth, government procurement, industrial projects, infrastructure programs, healthcare utilization, construction pipelines, energy investment and private-sector expansion.

Forecasting methods should match the available data rather than being selected for technological novelty. For stable historical series, exponential smoothing or ARIMA models may be adequate. Where demand is affected by multiple explanatory variables, regression, gradient boosting or random forests can be tested. Recurrent neural networks may be appropriate only when sufficient sequential data exist. Model performance should be compared using out-of-sample validation.

Scenario analysis is particularly important in Saudi Arabia because major industrial and infrastructure projects can create step changes in demand. A base case, accelerated-investment case and disruption case can provide decision-makers with a range rather than a single forecast.

5.3 Layer 3: Supplier-Ecosystem Analysis

Localization cannot be sustained without capable firms. Supplier mapping should identify current Saudi manufacturers, distributors that could move upstream into production, multinational firms with local operations, raw-material suppliers, engineering service providers, technology partners and research institutions. Each entity can be profiled using capacity, certifications, workforce, machinery, customer base, financial strength, location, expansion potential and technology readiness.

Network analysis adds an important relational perspective. Companies can be represented as nodes while supplier-buyer relationships are represented as edges. Centrality measures can reveal firms that occupy important positions across multiple value chains. Community-detection algorithms can reveal industrial clusters, and dependency analysis can identify nodes whose failure could affect several downstream firms.

Supplier scoring should distinguish reported capability from verified capability. Corporate websites and catalogues are useful discovery tools, but technical validation may require site visits, certifications, production records, customer references or third-party audits.

5.4 Layer 4: Value-Chain Localization Potential

Product-level localization should be decomposed into major value-adding stages such as raw-material preparation, intermediate processing, component manufacturing, assembly, testing, packaging, distribution and maintenance. The objective is to identify what share of a product's economic and technical value can realistically be performed in Saudi Arabia.

A Localizable Value Share can be estimated as the value of locally feasible activities divided by total product value. This prevents superficial localization from receiving the same score as deep localization. Final assembly may create visible domestic production but could leave most value embedded in imported components and intellectual property.

The value-chain layer should also identify enabling gaps. For example, a downstream manufacturing opportunity may be attractive only if specialty materials, testing facilities, tooling, regulatory certification or maintenance services are developed simultaneously. Localization programs can therefore be designed as ecosystem interventions rather than isolated factory projects.

5.5 Layer 5: Strategic Importance

Strategic importance introduces national-resilience considerations into the model. Products related to healthcare security, food security, energy systems, water, defense, critical infrastructure or essential

© SEP 2026 | IRE Journals | Volume 10 Issue 3 | ISSN: 2456-8880
DOI: https://doi.org/10.64388/IREV10I3-1723181

industrial processes may deserve priority even when commercial returns alone are moderate. This dimension ensures that the LOI is not merely a profitability ranking but a strategic industrial-development tool.

5.6 Layer 6: Economic-Impact Assessment

Economic-impact analysis estimates the potential national benefits of localization. Direct measures include domestic value added, investment, employment, local procurement and trade-balance effects. Indirect effects can be estimated through input-output relationships that capture upstream purchases and downstream activity. Induced effects may arise when wages generated by the project circulate through the wider economy. Technology transfer, export potential, SME participation and skill development should also be incorporated even where they are difficult to monetize precisely. A high-impact opportunity is one that not only replaces imports but creates an industrial platform with spillovers into adjacent sectors.

VI. LOCALIZATION OPPORTUNITY PRIORITIZATION

The six dimensions can be translated into a two-axis strategic matrix using market attractiveness and localization feasibility. This matrix supports a simple decision logic that is easy to communicate to senior stakeholders while preserving the more detailed LOI underneath.

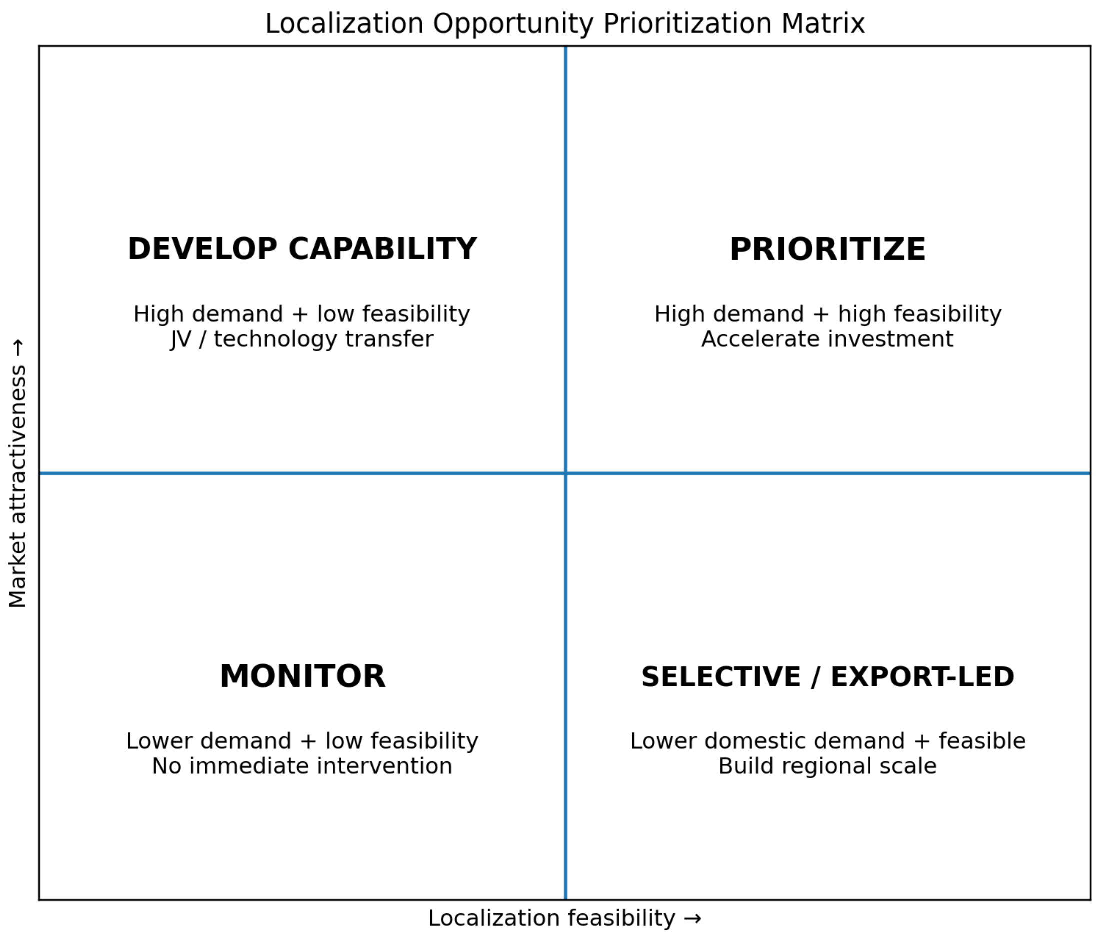

Figure 3. Localization Opportunity Prioritization Matrix. Source: Author-developed framework.

© SEP 2026 | IRE Journals | Volume 10 Issue 3 | ISSN: 2456-8880
DOI: https://doi.org/10.64388/IREV10I3-1723181

Table 3. Localization decision logic.

Market attractiveness	Feasibility	Classification	Recommended action
High	High	Immediate priority	Accelerate investment, incentives, supplier onboarding and investor matching
High	Low	Capability gap	Technology transfer, joint venture, workforce and supplier-development program
Low	High	Selective/export-led	Assess GCC/regional demand and export-oriented scale
Low	Low	Monitor	No immediate intervention; update as technology or demand changes

VII. ILLUSTRATIVE NATIONAL SCREENING PROCESS

Stage 1 - National trade screening. Screen imported products at HS6 or a more detailed level using minimum value, multi-year consistency, growth and origin concentration.
Stage 2 - Sector clustering. Group shortlisted products into coherent industrial families such as chemicals, pharmaceuticals, machinery, electrical equipment, food, automotive and construction materials.
Stage 3 - Demand validation. Link customs values with procurement, project and buyer-level evidence to distinguish recurring demand from one-time imports.
Stage 4 - Domestic capability mapping. Identify existing Saudi production and adjacent capabilities that could support expansion into shortlisted products.
Stage 5 - Value-chain assessment. Determine which production stages, inputs and services can realistically be localized and which gaps require technology or investment.
Stage 6 - Economic-impact modelling. Estimate value added, investment, employment, trade-balance effects and supplier spillovers.
Stage 7 - Strategic prioritization. Calculate LOI scores, test sensitivity to weights and assign opportunities to strategic categories.
Stage 8 - Investor and supplier matching. Translate analysis into action by matching priority opportunities with domestic investors, international technology owners and capable Saudi suppliers.

VIII. APPLICATION OF ARTIFICIAL INTELLIGENCE AND MACHINE LEARNING

8.1 Product clustering
Unsupervised learning can cluster thousands of products using import value, growth, concentration, volatility, complexity and domestic-supply variables. Clusters can distinguish high-value strategic imports, fast-growing intermediate goods, specialized capital equipment and mature consumer categories. Clustering is useful for exploration but should not replace sector expertise.

8.2 Supplier matching
Recommendation algorithms can compare manufacturing requirements with Saudi supplier profiles. Process capability, materials handled, machinery, certifications and workforce can be converted into feature vectors, enabling similarity scores that identify firms capable of moving into adjacent products.

8.3 Demand forecasting
Machine-learning models can estimate future demand where large, rich datasets exist. However, validation against simpler statistical baselines is essential. A model that is slightly more accurate but substantially less explainable may be less useful for industrial-policy decisions.

8.4 Natural-language processing
NLP can extract product, process and capability information from company websites, industrial catalogues, technical documents and investment

announcements. This can enrich structured supplier databases, but automated extraction should be verified before it influences high-value investment decisions.

8.5 Explainable AI

Explainability is especially important when public incentives or strategic designations depend on algorithmic outputs. Models should provide feature contributions, sensitivity analysis and auditable data lineage so decision-makers can understand why an opportunity is ranked highly.

IX. FINDINGS AND DISCUSSION

The framework produces five principal findings. First, localization is fundamentally a data-integration problem. Customs data show what is imported but not whether local firms can manufacture it. Supplier data show firms but not necessarily market demand. Procurement data validate buyers but may not capture future projects. The analytical advantage emerges when these sources are linked into a common product-company-value-chain architecture.

Second, product-level analysis is more actionable than broad sector analysis. Labels such as "healthcare manufacturing" or "automotive localization" are strategically useful but too aggregated for investment decisions. Within the same sector, a basic consumable, specialty chemical, diagnostic reagent and advanced electronic system have radically different capital, technology, regulatory and supplier requirements. Granular screening should therefore precede sector-level program design.

Third, supplier ecosystems are as important as market size. An isolated factory that imports nearly all equipment, components and technical services may generate limited domestic spillovers. A smaller project embedded in a network of local materials, tooling, maintenance, logistics and engineering providers can generate deeper value addition and stronger resilience. Localization policy should therefore measure depth as well as headline production.

Fourth, digital transformation has a dual function. It strengthens the competitiveness of localized manufacturers while also improving the information available to industrial decision-makers. Digital procurement, supplier-performance systems, production data and logistics visibility create a richer evidence base for future localization decisions.

Fifth, localization rankings must remain dynamic. Technology costs change, new factories open, demand forecasts are revised, trade relationships evolve and product substitution occurs. A one-time industrial study becomes outdated. The proposed system should therefore operate as a continuously refreshed decision-support capability with periodic re-ranking and expert review.

X. ALIGNMENT WITH SAUDI VISION 2030

10.1 Economic diversification

Prioritizing competitive domestic production expands non-oil economic activity and reduces reliance on a narrow set of revenue sources. The framework supports diversification by directing attention to products and value chains with measurable demand and feasible domestic capability.

10.2 Local-content development

The LOI can support local-content strategy by identifying categories where domestic value capture can increase through materials, manufacturing, engineering, testing, maintenance or services rather than focusing only on final assembly.

10.3 Investment attraction

Investors often face uncertainty regarding market size, competitors, buyers and supply-chain readiness. A data-driven localization platform can convert fragmented evidence into bankable opportunity profiles containing demand history, forecast growth, supplier maps and value-chain gaps.

10.4 SME development

Large localized industries create opportunities for SMEs in components, packaging, tooling, maintenance, logistics, software and technical services. Supplier-network analytics can identify where smaller Saudi firms can participate and where capability-development programs should be targeted.

10.5 Supply-chain resilience
Strategic products with concentrated foreign sourcing can be identified and prioritized for domestic alternatives, regional sourcing, inventory measures or technology partnerships. The framework therefore supports resilience even when full localization is not the preferred response.

10.6 Digital and human capability
Building and operating the proposed platform requires data engineers, industrial economists, sector experts, data scientists and policy professionals. It therefore reinforces the Vision 2030 objective of developing higher-value human and digital capabilities.

XI. IMPLEMENTATION ARCHITECTURE AND GOVERNANCE

11.1 Four-Layer Technology Architecture

Table 4. Proposed implementation architecture.

Layer	Main components	Primary output
1. Data foundation	Trade, production, procurement, company, supplier, investment and logistics data	Trusted integrated data assets
2. Analytics engine	SQL, Python, statistical modelling, ML, network analysis	Pattern detection, forecasting and scoring
3. Decision models	Dependency, demand, supplier, value-chain, strategic and impact models	LOI and scenario outputs
4. User interface	Power BI or equivalent dashboards and opportunity profiles	Executive and analyst decision support

Data governance should assign clear ownership for each source and define standards for completeness, timeliness, classification, security and access. Analytical governance should document model assumptions, weights and data lineage. Institutional governance should establish how industrial, trade, investment, procurement and statistical organizations exchange information and resolve disagreements.

Model outputs should remain explainable. A senior decision-maker should be able to see whether an opportunity is ranked highly because of demand, import concentration, supplier readiness, strategic criticality or another factor. This transparency is particularly important when public incentives, land, financing or regulatory support are involved.

Cybersecurity and privacy controls are also necessary because supplier and procurement information can contain commercially sensitive data. Access should be role-based, data-sharing agreements should be explicit, and the system should comply with relevant Saudi data-management, cybersecurity and privacy requirements.

XII. CHALLENGES AND RISKS

Data fragmentation: Different organizations may use separate systems, classifications and identifiers. Integration requires common taxonomies, metadata standards and entity-resolution rules.

Data quality: Incorrect HS classification, duplicated company records and inconsistent quantities can distort opportunity rankings. Automated quality checks and expert review are necessary.

Capability verification: Company claims may not equal actual production capacity. Supplier readiness should be verified using certifications, equipment, production evidence and technical assessment.

Skills shortage: The platform requires professionals who combine data skills with industrial, commercial

© SEP 2026 | IRE Journals | Volume 10 Issue 3 | ISSN: 2456-8880
DOI: https://doi.org/10.64388/IREV10I3-1723181

and policy understanding. Purely technical teams may miss sector realities, while purely commercial teams may underuse analytical methods.

Algorithmic bias: Historical procurement may reflect legacy sourcing patterns rather than future priorities. Models should combine historical evidence with strategic and forward-looking variables.

Commercial viability: Not every strategically desirable product can be manufactured competitively. Incentives should not permanently compensate for structurally uncompetitive economics.

Change management: Organizations may resist common data standards or transparent prioritization. Leadership sponsorship and clear governance are therefore as important as technology.

XIII. POLICY AND MANAGERIAL RECOMMENDATIONS

1. Develop a unified localization data platform. Integrate trade, industrial, procurement, supplier and project data around common product and company identifiers.
2. Prioritize at product level. Use detailed product codes for initial screening before grouping opportunities into sector programs.
3. Adopt multi-criteria scoring. Do not use import value as the sole criterion; include demand, feasibility, strategic importance and economic impact.
4. Build verified supplier profiles. Maintain continuously updated capability records and distinguish verified manufacturing from distribution or assembly.
5. Integrate project pipelines. Use planned industrial, infrastructure and government investment to improve demand forecasting.
6. Use scenarios and sensitivity analysis. Test how rankings change under different growth, disruption and policy assumptions.
7. Measure post-investment outcomes. Track whether localization projects actually deliver value added, employment, domestic procurement, exports and resilience.
8. Link analytics with investor development. Convert high-ranked opportunities into structured investment cases and match them with Saudi and international investors.
9. Create feedback loops. Use real project outcomes to recalibrate LOI weights and improve future predictions.

XIV. RESEARCH IMPLICATIONS

14.1 Theoretical Implications
The proposed framework advances localization research by combining trade dependence, market attractiveness, industrial capability and strategic value in a single analytical model. Traditional import-substitution perspectives tend to emphasize trade values and protection mechanisms. The present framework instead views localization as a capability-building problem in which data, suppliers, technology and value-chain depth determine whether domestic production can become sustainable.

14.2 Managerial Implications
Large Saudi companies can adapt the framework to their own procurement portfolios. Category-level LOI scores can distinguish between products that should remain globally sourced, categories suitable for Saudi supplier development, categories that require international joint ventures and categories capable of deep local production. This allows local-content programs to become more evidence based and helps procurement teams focus supplier-development resources where the opportunity is largest.

14.3 Investor Implications
For investors, data-driven localization reverses the conventional sequence of manufacturing investment. Instead of beginning with a product idea and then searching for a market, the investor begins with a validated market gap, tests localization feasibility, identifies suppliers and only then develops the investment case. This process can reduce market-entry uncertainty and help industrial-development organizations present more credible opportunities.

XV. LIMITATIONS AND FUTURE RESEARCH

The study is conceptual and does not claim to report results from a complete confidential Saudi customs, procurement and supplier dataset. This is an important limitation because empirical application may reveal sector-specific relationships not captured in a generic framework.

Localization feasibility also varies substantially by sector. Pharmaceuticals, food, chemicals, electronics and industrial machinery require different technical and regulatory criteria. Future research should therefore develop calibrated sector-specific versions of the LOI.

The weighting system requires empirical validation. Expert elicitation, analytic hierarchy process methods, historical project outcomes and sensitivity analysis can be used to determine appropriate weights. Future studies should test whether high-LOI opportunities are more likely to produce successful industrial projects.

Geospatial analysis represents another important extension. Industrial opportunities can be matched with regions based on ports, industrial cities, customers, raw materials, utilities, workforce and logistics. This would allow the model to answer not only what should be localized but where.

Finally, supplier-network research can examine whether ecosystem density, centrality or diversity predicts localization sustainability. Such analysis would deepen understanding of how industrial clusters contribute to long-term competitiveness.

XVI. CONCLUSION

Saudi Arabia's industrial transformation under Vision 2030 requires increasingly sophisticated methods for determining where localization can create the greatest sustainable value. The scale of merchandise trade, diversity of industrial demand and complexity of modern value chains make purely qualitative or single-variable approaches insufficient. Big data analytics provides a practical mechanism for transforming fragmented information into industrial intelligence.

This paper developed a framework built around six dimensions: import dependency, market attractiveness, supplier capability, value-chain localization potential, strategic importance and economic impact. The framework's central conclusion is that high import value should be treated as the beginning, not the conclusion, of localization analysis. A viable opportunity must also have sufficient demand, an achievable technology and supplier pathway, meaningful local value capture and a clear economic or strategic rationale.

Integration of customs data, production records, supplier databases, procurement demand, investment pipelines and macroeconomic indicators can allow opportunities to be screened at a scale impossible through manual analysis alone. SQL and Python can support data preparation and modelling, business-intelligence tools can provide decision transparency, and machine-learning methods can improve clustering, demand forecasting and supplier matching where appropriate.

The approach aligns closely with Vision 2030 objectives relating to non-oil growth, industrial development, local content, investment attraction, SME participation, digital capability and resilient supply chains. Its value is not limited to government. Large companies can use the same principles to prioritize procurement localization, while investors can use the outputs to identify market gaps backed by evidence.

Ultimately, successful localization is not simply the substitution of foreign products with domestic products. It is the creation of competitive and resilient industrial ecosystems that can sustain quality, technology, productivity and market demand. A continuously updated data-driven localization platform can help Saudi Arabia move from broad industrial ambition to specific, measurable and investable opportunities, strengthening the Kingdom's capacity to capture value from its domestic market while developing globally competitive non-oil industries.

REFERENCES

[1] A. Aljuaid, "Integrating Industry 4.0 for sustainable localized manufacturing to enhance logistics performance," *Sustainability*, vol. 16, no. 12, Art. no. 5096, 2024.

[2] General Authority for Statistics (GASTAT), *International Merchandise Trade Statistics: May 2025*. Riyadh, Saudi Arabia, 2025.

[3] General Authority for Statistics (GASTAT), *International Merchandise Trade Statistics: June and Q2 2025*. Riyadh, Saudi Arabia, 2025.

[4] General Authority for Statistics (GASTAT), *Methodology and Quality Report for International Trade in Goods Statistics*. Riyadh, Saudi Arabia, 2025.

[5] Saudi Vision 2030, *Saudi Vision 2030: Overview*. Kingdom of Saudi Arabia, 2026.

[6] Saudi Vision 2030, *National Industrial Development and Logistics Program*. Kingdom of Saudi Arabia, 2026.

[7] Saudi Vision 2030, *National Industrial Strategy*. Kingdom of Saudi Arabia, 2025.

[8] Saudi Vision 2030, *Saudi Arabia's Supply Chain Playbook*. Kingdom of Saudi Arabia, 2026.

[9] Saudi Vision 2030, *Public Investment Fund Program 2021–2025 Delivery Plan*. Kingdom of Saudi Arabia, 2021.

[10] A. M. Asiri *et al*., "The integration of sustainable technology and big data analytics in Saudi Arabian SMEs," *Sustainability*, vol. 16, no. 8, Art. no. 3209, 2024.

[11] S. Alorfi *et al*., "Barriers to the adoption of big data analytics in Saudi organizations," *Systems*, vol. 13, no. 4, Art. no. 250, 2025.

[12] Tripathi *et al*., "Digital transformation readiness and business performance among SMEs in Saudi Arabia," *Sustainability*, vol. 16, no. 9, Art. no. 3831, 2024.

[13] Mutambik, "Digital transformation, sustainability and competitiveness in Saudi Arabia," *Sustainability*, vol. 16, no. 10, Art. no. 4310, 2024.

[14] T. H. Alaskar, K. Mezghani, and A. K. Alsadi, "Big data analytics adoption in supply chain management under competitive pressure: Evidence from Saudi Arabia," *Journal of Decision Systems*, 2020.

[15] A. K. Alsadi, T. H. Alaskar, and K. Mezghani, "Adoption of big data analytics in supply chain management: Organizational factors and supply-chain connectivity," *International Journal of Information Systems and Supply Chain Management*, vol. 14, no. 2, pp. 88–107, 2021, doi: 10.4018/IJISSCM.2021040105. IGI Global

[16] N. F. Alogaiel *et al*., "An assessment of the quality of open government data in Saudi Arabia," *IEEE Access*, 2023.

[17] Z. A. Khan *et al*., "Industrial sustainability and manufacturing transition in the Kingdom of Saudi Arabia," *IEEE Access*, 2023.

[18] Zaki *et al*., "Sustainability-oriented innovation in Saudi SMEs," *Administrative Sciences*, vol. 15, no. 2, Art. no. 59, 2025.

[19] M. Christopher and H. Peck, "Building the resilient supply chain," *The International Journal of Logistics Management*, vol. 15, no. 2, pp. 1–14, 2004, doi: 10.1108/09574090410700275. Emerald

[20] D. Ivanov and A. Dolgui, "Viability of intertwined supply networks: Extending supply-chain resilience angles towards survivability," *International Journal of Production Research*, vol. 58, no. 10, pp. 2904–2915, 2020.

[21] S. F. Wamba, A. Gunasekaran, S. Akter, S. J. Ren, R. Dubey, and S. J. Childe, "Big data analytics and firm performance: Effects of dynamic capabilities," *Journal of Business Research*, vol. 70, pp. 356–365, 2017, doi: 10.1016/j.jbusres.2016.08.009. ScienceDirect

[22] R. Dubey, A. Gunasekaran, S. J. Childe, C. Blome, and T. Papadopoulos, "Big data and predictive analytics and manufacturing performance," *British Journal of Management*, vol. 30, no. 2, pp. 341–361, 2019, doi: 10.1111/1467-8551.12355. Wiley

[23] F. Kache and S. Seuring, "Challenges and opportunities of digital information at the intersection of big data analytics and supply chain management," *International Journal of Operations & Production Management*, vol. 37, no. 1, pp. 10–36, 2017, doi: 10.1108/IJOPM-02-2015-0078. Emerald

[24] M. M. Queiroz, R. Telles, and S. H. Bonilla, "Blockchain and supply chain management integration: A systematic review," *Supply Chain Management*, vol. 25, no. 2, pp. 241–254, 2020.

[25] D. Ivanov, "Supply Chain Viability and the COVID-19 pandemic: A conceptual and formal generalisation of four major adaptation strategies," *International Journal of Production Research*, vol. 59, no. 12, pp. 3535–3552, 2021.

[26] World Bank, *World Development Report 2020: Trading for Development in the Age of Global Value Chains*. Washington, DC, USA: World Bank, 2020. World Bank

[27] OECD, "Global value chains: Efficiency and risks in the context of COVID-19," *OECD Policy Responses to Coronavirus (COVID-19)*, 2021, doi: 10.1787/67c75fdc-en. OECD

[28] UNIDO, *Industrial Development Report 2022: The Future of Industrialization in a Post-Pandemic World*. Vienna, Austria: UNIDO, 2022.

[29] M. E. Porter, *The Competitive Advantage of Nations*. New York, NY, USA: Free Press, 1990.

[30] Humphrey and H. Schmitz, "How does insertion in global value chains affect upgrading in industrial clusters?," *Regional Studies*, vol. 36, no. 9, pp. 1017–1027, 2002, doi: 10.1080/0034340022000022198. Taylor & Francis