

Data and AI for Tax Revenue Forecasting Innovation

Supporting Fiscal Policy Decision-Making in Korea

2025. 9. 26.

Yeonhee Lee

Director, Fiscal and Economic Policy Intelligence Research Center
Electronics and Telecommunications Research Institute (ETRI)

Outline

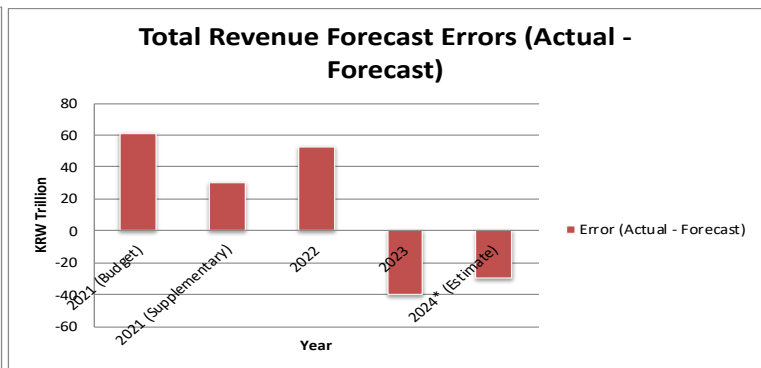
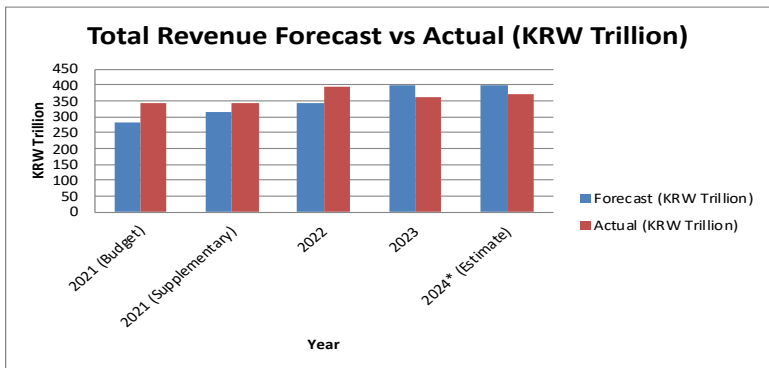
- Why Tax Revenue Forecasting Matters
- Why Corporate Tax Forecasting Matters
- AI Evolution and Roles in Fiscal Policy
- Global Practices in Tax Administration Digitalization
- Korea's Fiscal AI Initiative & Fiscal World Model
- Results: Accuracy, Timeliness, Corporate Tax Forecasting
- Policy Implications and Closing

Korea's Fiscal AI Initiative

- Conducted by ETRI, Korea's national ICT research institute (MSIT)
- Commissioned by Ministry of Economy and Finance (MOEF)
- Focus: AI-driven tax forecasting, digital twins, ABM models
- Goal: Enhance accuracy, agility, and trust in fiscal policy

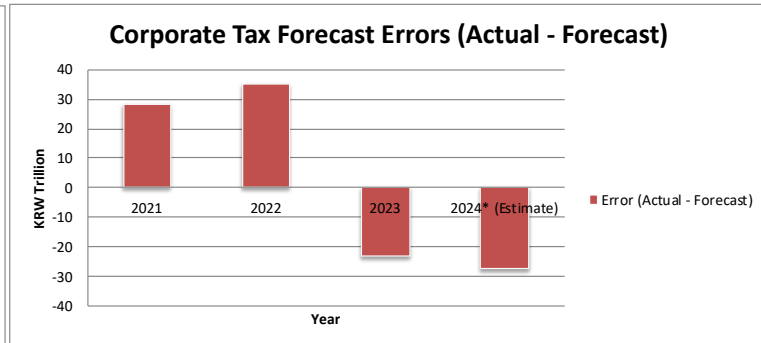
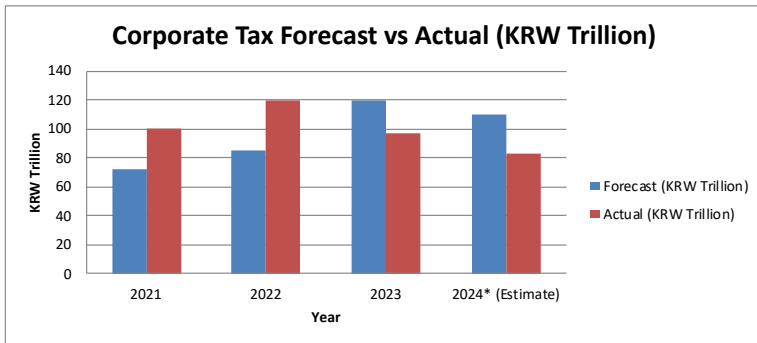
Why Tax Revenue Forecasting Matters

- Fiscal planning depends on reliable revenue forecasts
- Korea's recent large forecast errors (2021–2023: tens of trillions KRW) undermined fiscal stability.
- Similar issues across OECD countries
- Need for real-time, explainable, and adaptive forecasting models



Why Corporate Tax Forecasting Matters

- Corporate tax is the most volatile and error-prone component of revenue
- Korea's corporate tax forecast errors (2021–2023) drove a significant share of total revenue gaps.
- Similar corporate tax forecast challenges exist across OECD countries
- Need for specialized models with micro-level and timely firm data



The Evolution of AI: From Rules to Domain-Specific Intelligence



1st Generation: Rule-Based Systems (1980s–1990s)

- IF–THEN rules, limited domains
- Difficult to adapt to changes

2nd Generation: Machine Learning & Data-Driven (2000s–2010s)

- Statistical learning, pattern recognition
- Limited performance without large-scale data

3rd Generation: Deep Learning & General AI (2010s–2020s)

- Breakthroughs in image, speech, and text
- Emergence of large-scale LLMs such as GPT
- Limitations: accuracy, explainability, lack of domain knowledge

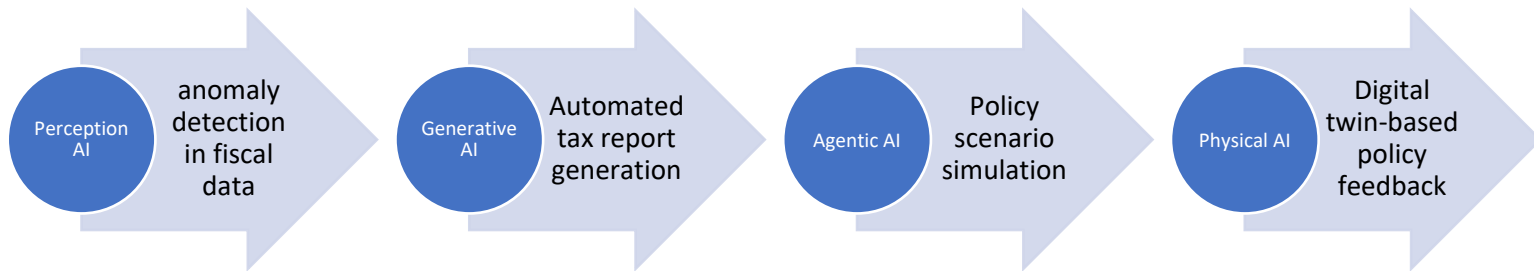
4th Generation: Domain-Specific AI (Present)

- Tailored to industry/policy domains
- Incorporates regulation, data, and context → enhanced reliability
- Examples: Indonesia's *KemenkeuGPT*, *TaxAI* research

AI Evolution in Fiscal Policy(Function/Role)

- **Perception AI > Generative AI > Agentic AI > Physical AI**

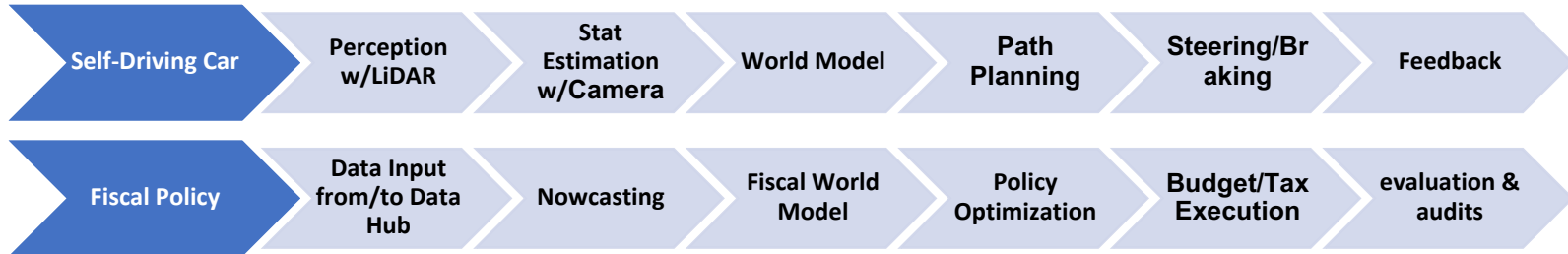
- Perception AI – anomaly detection in fiscal data
- Generative AI – auto-generation of reports & scenarios
- Agentic AI – policy simulation & decision support
- Physical AI – integration with fiscal digital twins



Self-Driving vs Fiscal Policy Architecture

- Pipelines for Policy Autonomy

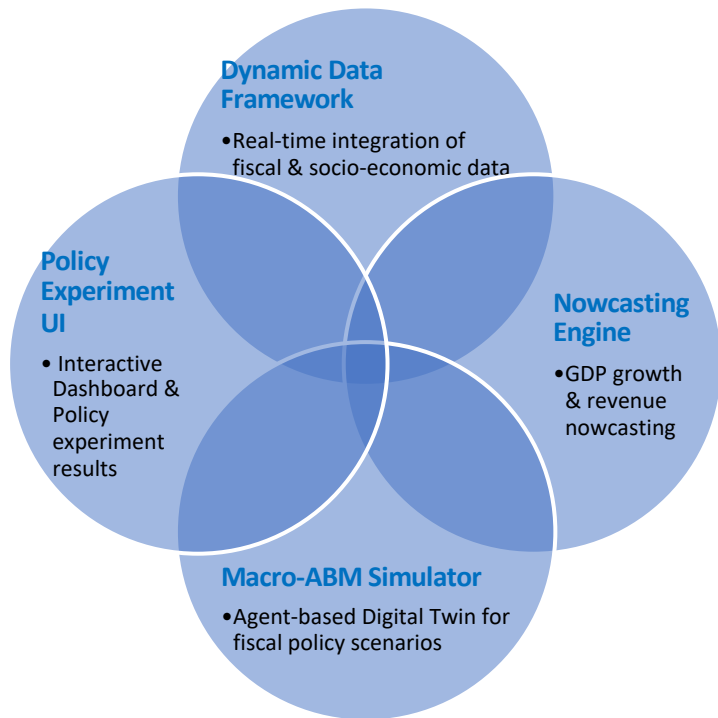
“Both require a cycle of perception, forecasting, decision, and execution.”



Global Practices in Tax Administration Digitalization

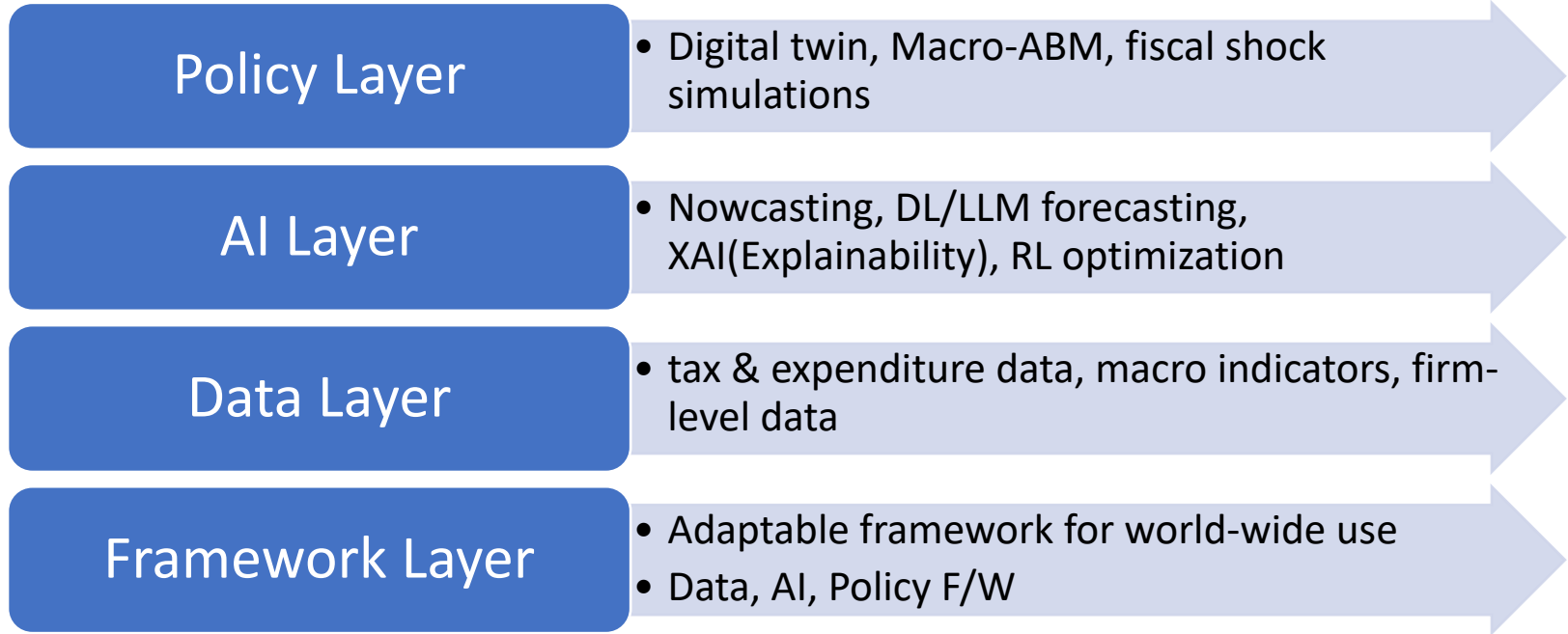
- Real-time & High-Frequency Data: EU (ViDA), Spain (SII), Australia (STP)
 - Nowcasting enabled by real-time reporting
- Invoice & Data Quality: Italy (SdI), Japan (QIS), Canada (E-filing)
 - Standardization reduces VAT gap & improves accuracy
- Governance & Institutional Design: UK (OBR/HMRC), US (Treasury/CBO/JCT)
 - Independent forecasting & API integration
- Strong Baseline Systems: Korea (NTS)
 - Mandatory e-invoicing, next-day submission
 - Strong basis for VAT & corporate tax forecasting
- Summary:
 - Timeliness and data quality are critical for reliable tax revenue forecasting
 - Countries are converging toward real-time reporting, invoice standardization, and strong governance frameworks
 - These practices strengthen Nowcasting and forecasting reliability

Korea's AI Fiscal Framework: Simplified Architecture



- Fiscal Digital Twin: Fiscal World Model = Core Engine (Macro-ABM, SFC, AI)
 - Surrounding Components: Data pipelines, dashboards, policy execution systems
- 👉 **Key Challenge: Who builds a trusted Fiscal World Model first?**

Korea's AI Fiscal Framework



Simple vs Complex



Why Do We Need a World Model?

- “Autonomous driving cannot be achieved with simple rules; Likewise, policy cannot rely on averages or regressions alone.”
- “In complex environments, a World Model is essential — That is why Digital Twins, Macro-ABM, and AI are needed for public policy.”

From Self-Driving Cars to Self-Driving Policy

Category	Road World Model	Fiscal World Model (Macro-ABM DT)
Domain	Physical traffic system	Socio-economic & fiscal system
Agents	Vehicles, pedestrians, traffic signals	Government, firms, households, international sector
Input Data	Sensors, cameras, LiDAR, traffic signals	Revenue & expenditure data, macro indicators, administrative guidelines
Outputs (Objectives)	Route planning, safety, efficient traffic flow	Revenue forecasting, policy effects, fiscal stability
Complexity Features	Nonlinear interactions, accident/congestion propagation	Nonlinear policy effects, business cycle & revenue fluctuations
Use Case (Purpose)	Real-time decision-making (ultra-short term)	Mid- to long-term policy & fiscal decision support

Policy Impact – So What?

- Accuracy – Reduce forecast errors, restore trust
- Agility – Simulate shocks, support timely response
- Efficiency – Better budget/expenditure monitoring
- Trust – Transparent, explainable models

Practical Results – Highlights

- **Accuracy** ↑ – Forecast errors reduced by ~30% with AI models
- **Timeliness** ↑ – Forecast update cycle shortened (Quarterly → Monthly → Weekly)
- **Corporate Tax Focus** – AI-driven corporate earnings forecasting shows clear gains (MAE ↓, MAPE ↓)
- **Trust** ↑ – More transparent and explainable forecasts

Traditional Macroeconomic Models vs Korea's Fiscal AI Approach

OECD/IMF & Traditional Models

- Focus on GDP, inflation, employment, fiscal balance
- Corporate profits only as aggregate sector variable
- No direct firm-level operating profit forecast
- Regression & time-series based (ARIMA, VAR, DSGE, SFC)

Korea's Fiscal AI Initiative

- Focus on Firm-level operating profit forecast
- LLMs + deep learning for corporate earnings
- Integrates financials, news, market data
- Links profits → tax revenue → fiscal forecasts
- Digital Twin = Fiscal World Model

We complement traditional macro models by linking firm-level AI forecasts to fiscal outcomes.

Baseline vs AI – Corporate Earnings Forecasting

Baseline Approaches

- Analyst Consensus
- Time-series (ARIMA, VAR)
- Machine Learning (Random Forest, Boosting)
- Deep Learning (LSTM, GRU)

Performance:

- MAE: 9,000–12,000 (100M KRW)
- MAPE: 35–45%

AI Approach (Fiscal AI Initiative)

- LLM-based Forecasting (LLM + News Summarization Agent)
- Data: Financials + News + Market + ...
- Linked to Corporate Tax Forecasts

Performance:

- MAE: ~6,100 (100M KRW), ↓ ~2,400 vs Consensus
- MAPE: ~11%, ↓ ~14%p vs Consensus
- Consistent improvement across 4 years

AI + News Summarization Agent with LLM outperforms consensus forecasts: ≈2,400 (100M KRW) lower MAE, ≈14%p lower MAPE, consistently across 4 years.

Policy Implication

- **Data First: Micro & Timely Data Infrastructure**
 - High-frequency, standardized, high-quality data
 - **High-Fidelity Models: Macro-ABM, Digital Twins**
 - Move beyond regressions → simulate complex policy shocks
 - **Governance & Trust – Transparency, Accountability**
 - AI as a support tool, not black-box decision-maker
- ❖ Supporting:
- AI & Digital Tools (nowcasting, explainability, UQ)
 - Capacity & Skills (interdisciplinary training, collaboration)

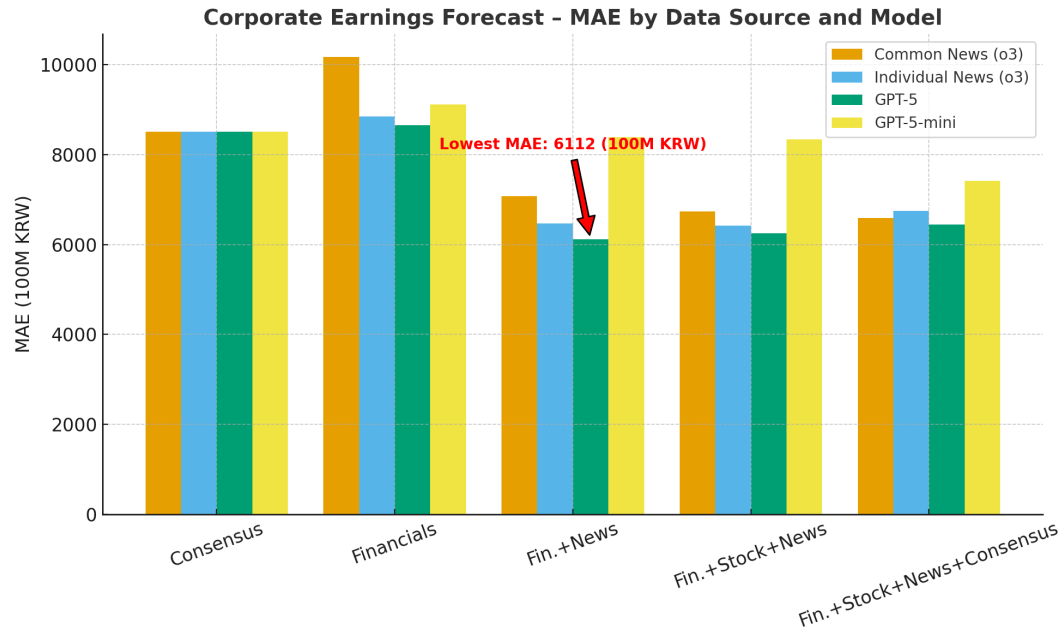
Closing



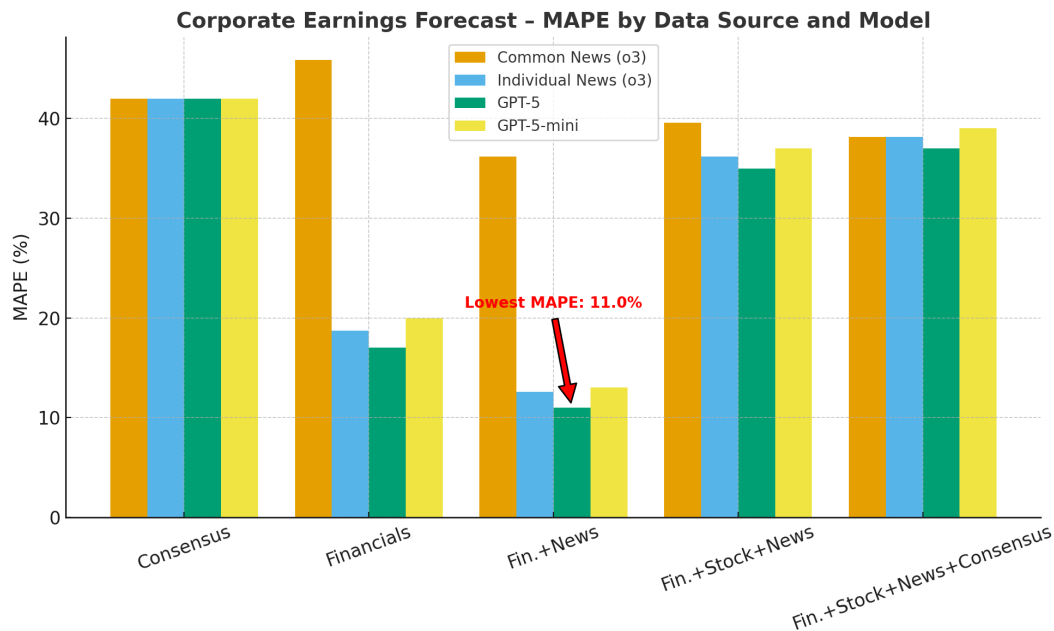
The convergence of Data + Models + AI builds a trusted Fiscal World Model for smarter policy.

Appendix

Appendix – Corporate Earnings Forecasting (MAE)



Appendix – Corporate Earnings Forecasting (MAPE)



Appendix – From Sensors to Decisions: A Layered Architecture

Perception Layer (Sensors / Data In fra)

- **Road:** LiDAR, cameras → perception of road environment
- **Fiscal:** Real-time streams of revenue, expenditure, economic, and social data

State Estimation

- **Road:** Current position, speed, and traffic conditions
- **Fiscal: Current state of the economy and public finances (Nowcasting)**

World Model / Simulation

- **Road:** Road World Model → route and traffic flow prediction
- **Fiscal:** Macro-ABM, DSGE, System Dynamics, and SFC-based fiscal simulations

Planning & Decision

- **Road:** Optimal route selection, lane changes, obstacle avoidance
- **Fiscal:** Optimization across revenue stability, growth, and equity objectives

Actuation & Execution

- **Road:** Steering, acceleration, braking
- **Fiscal:** Tax law changes, budget allocation and execution, subsidies and taxation policies

Feedback & Learning

- **Road:** Driving logs → continuous model updates
- **Fiscal:** Policy outcomes and audit results → continuous model updates

Appendix – Global Practices in Tax Administration Digitalization: Detailed

Country / Region	Approach & Framework	Digitalization / AI Initiatives	Key Implications
EU (ViDA)	Cross-border e-Invoicing & phased mandatory real-time reporting	VAT fraud reduction, enhanced forecast data	Strengthening Nowcasting infrastructure through common standards
Italy	Mandatory B2B e-Invoicing (SdI)	Significant reduction in VAT gap (official statistics)	Invoice data quality improves forecasting accuracy
Spain	SII (real-time VAT ledger) system	High-frequency, near real-time reporting	Tax revenue Nowcasting enabled by granular data
UK (OBR/HMRC)	OBR independent forecasting + HMRC tax item models	Making Tax Digital, API integration	Case of governance linked with digitalization
USA (Treasury/CBO/JCT)	Split institutional roles (policy, forecasting, revenue)	IRS Direct File pilot (e-filing)	Standardization & improved in-year forecast accuracy
Australia (ATO)	Traditional models + real-time payroll data	STP (Single Touch Payroll)	Payroll/withholding-based Nowcasting foundation
Canada (CRA)	Fiscal-macroeconomic models, expanded automation	Enhanced e-filing services, DST introduction	Digitalized reporting → improved data quality
Korea (NTS)	Tax-item models, linked with macro indicators	Mandatory e-invoicing (2011~), next-day submission	Strong basis for VAT & corporate tax forecasting
Japan (NTA)	Combined macro-tax models	Qualified Invoice System (2023~)	Invoice standardization improves consumption tax accuracy