1 Data

This section, provides an overview of data to be used for our prediction model. The target variable is R&D expenditure, known as Gross Domestic Expenditure on Research and Development (GERD), which is published on a yearly basis by OECD. We use mixed-frequency data, consisting of macroeconomics variables and GT data, to generate nowcasting values for yearly R&D, and offer the more granular ones, afterwards.

1.1 Gross Domestic Expenditure on Research and Development (GERD)

The target variable of the study is the GERD, which encapsulates the total expenditure (both current and capital) on R&D activities conducted by all resident organizations, including companies, research institutes, universities, and government laboratories within a nation's territory [OECD [2023]]. According to the OECD, this measurement includes R&D funded from external sources, but excludes domestic funds designated for R&D activities outside the domestic economy [OECD [2015]].

1.2 Google Trends (GT)

To gain better insights into R&D expenditure and its potential drivers, we established an ecosystem of stakeholders associated with R&D spending. This ecosystem comprises entities that either directly partake in R&D activities or play a supporting role in the broader R&D landscape. For each stakeholder, we used the expert knowledge to identify the specific search terms based on the likelihood that they would be searched in relation to R&D activities. This was executed to build a high-dimensional set of predictors grounded in monthly GT data.

Table 1: Initial search terms (GT) to build network of stakeholders

Stakeholder	**Keywords/Topics/Categories**
Firms/Companies	R&D Expenditure, Product Development, Technology Innovation, Patent Application, Tech Research, Pharma Research, New Drug Application, Research Grants, Intellectual Property.
Venture Capitalists (VCs)	Startup Funding, Technology Startups, Pharma Startups, VC Investment in R&D, Innovation Investment, Return on Investment, Exit Strategy, Seed Funding, Angel Investment.
Banks and Financial Institutions	Business Loans, R&D Loans, Investment Banking, Corporate Finance, Financial Risk, Credit Assessment, Interest Rate, Loan Application, Credit Score.
Universities and Research Institutions	Academic Research, Collaboration with Industry, Research Funding, University Patents, Postgraduate Studies, Doctoral Research, Research Publication, Research Grant.
Government agencies	R&D Policy, Research Funding, Government Grants, Innovation Policy, Public-Private Partnership, Tax Incentives for R&D, Technology Transfer, Patent Law, Economic Development.
R&D Employees	Research Methods, Data Analysis, Patent Filing, Lab Equipment, Scientific Journal, Professional Development, Research Ethics, Project Management, Collaboration Tools.
Tax Authorities	R&D Tax Credit, Tax Incentives, Tax Deduction, Tax Filing, Corporate Tax, Tax Law, Tax Consultancy.
Consulting Firms	Business Strategy, Market Analysis, Risk Assessment, Business Growth, Innovation Strategy, Portfolio Management, Project Planning, Financial Modeling.
Innovation Hubs/Incubators	Startup Incubation, Innovation Hub, Technology Park, Business Accelerator, Entrepreneurship, Mentorship, Networking, Business Pitch, Startup Ecosystem.
Patent Attorneys	Patent registration, Intellectual property rights, Patent law, Technology patents, Patent disputes.
Tax Consultants/Accountants	R&D tax credits, Corporate tax, Business expenses, Tax deductions, Tax advice for R&D, Accounting for R&D.

While the initial ecosystem developed for this study encompassed a broader range of search terms, detailed in table 1, table 2, presents the finalized selection of stakeholders and their corresponding search terms, which have been carefully selected for use in the input vector of our model.

Table 2: Stakeholders and their respective search terms for R&D expenditure

Stakeholder	**Search Terms**
Businesses	R&D Expenditure, Product Development
Consulting Firms	Innovation Management, Innovation Strategy
Government Agencies	Government Grants, Research Funding
Innovation Hubs	Startup Incubation, Technology Park
Patent Attorneys	Patent Attorney, Patent Registration
R&D Employees	R&D Jobs
Research Institutions	Collaboration with Industry, Research Grant
Tax Authorities	R&D Tax Credit
Venture Capitalists (VCs)	VC Investment, Startup Funding

1.3 Macroeconomic Variables

Additionally, the study also incorporated macroeconomic variables as potential predictors. These include:

- Gross Domestic Product per Capita: Reflects the wealth and economic strength of a nation.
- Unemployment Rate: Indicates the health of the labor market and broader economic conditions.
- Population: Used as a scale variable, representing the size of a nation.
- Inflation Rate: Represents the rate of price growth and can indicate the state of the economy.
- Export Volume & Import Volume: Chosen as indicators of a nation's competitiveness and its engagement with the global economy.

The data for these variables were sourced from the International Monetary Fund (IMF)'s World Economic Outlook Database International Monetary Fund [2023]. These variables provide a comprehensive overview of a nation's economic state, capturing elements of wealth, economic conditions, and competitiveness, which could all influence R&D expenditure.

References

OECD. Gross domestic spending onR&D(indicator), 2023. URL 10.1787/d8b068b4-en. Accessed on 27 October 2023.

OECD. *Measuring R&D: Methodologies and procedures*. OECD Publishing, Paris, 2015. doi:10.1787/9789264239012-8-en. URL 10.1787/9789264239012-8-en.

International Monetary Fund. World economic outlook database. https://www.imf.org/en/Publications/WEO/weo-database/2023/April/download-entire-database, April 2023. Accessed on 20 June 2023.