**Data Dive for Development Hackathon 2025**

**Participant Handbook**

**Theme:** Leveraging Data Science to Analyze and Enhance Development Cooperation

**Challenge:** UNDP Seoul Policy Centre is looking for **functional prototypes** that extract, analyze, and visualize insights from Official Development Assistance (ODA)

**Target participants:** Data scientists, developers, UI/UX designers, researchers, international development practitioners, tech enthusiasts; students and professionals welcome to apply.

**Eligibility:** This challenge is open to participants aged 19-34. All participants must reside in the Republic of Korea (ROK) at the time of application and for the duration of the agreement.

**Team size:** 2-4 members

**Awards:**

|  |  |  |
| --- | --- | --- |
| **Awardees** | **Agreement Amount** | **Remarks** |
| 3 teams | KRW 6,000,000 for each team | UNDP Data Dive for Development Hackathon Award |

**Key Dates:**

|  |  |  |
| --- | --- | --- |
| **Event** | **Time Period** | **Details** |
| **Information Session** | June 30 | Explanation of the hackathon challenge and Q&A opportunity for participants (online) |
| **Submission Period** | July 28 – August 3 | Submission of proposals |
| **Winners Announcement** | September 3 | Announcement of selected winners |
| **Orientation for Selected Teams** | W/C September 8 | Selected teams will attend orientation session to discuss further plans to develop their prototype |
| **Further Prototype Development** | W/C September 15 – W/C November 10  | Teams will continue their project development with guidance by UNDP |
| **Final Presentation** | W/C November 17 | Teams will share their final prototypes and research findings |

# Challenge Owner

As the lead United Nations agency on international development, UNDP works in 170 countries and territories to eradicate poverty and reduce inequality. We help countries develop policies, leadership skills, partnerships and institutional capabilities to achieve the Sustainable Development Goals. Our work is centred around six core development areas, known as our signature solutions: poverty and inequality, governance, resilience, environment, energy and gender equality.

The UNDP Seoul Policy Centre (USPC) is a facilitator of innovative development cooperation to catalyse the achievement of the Sustainable Development Goals. Through its SDG Partnerships programme and other South-South and Triangular Cooperation initiatives, the Centre supports countries by sharing innovative, tested-and-proven practices and policy tools of the Republic of Korea. The Centre's programmatic focus comprises i) Development Cooperation Approaches and Modalities; ii) Governance and Gender; iii) Green Recovery and Transition; and iv) Private Sector Engagement and Development.

# Background – What We are Looking For

This hackathon is a unique competition inviting participants to develop **innovative microservices** to derive insights from data on Official Development Assistance (ODA). UNDP Seoul Policy Centre will select three winning teams, awarding them a two-month agreement to refine their initial prototypes.

USPC is looking for **prototype solutions** with potential to be scaled to support both our internal processes and work on the ground. By developing microservices that effectively combine a comprehensive ODA database with complementary datasets, we aim to uncover hidden ODA patterns, predict future trends, and better understand the complex relationships that shape development cooperation. The ultimate aim is to use these insights to inform better decision-making and effectiveness across our Centre operations.

This hackathon represents an opportunity to apply cutting-edge data science techniques to transform data into actionable intelligence that can inform better policymaking and more effective development cooperation.

# Potential Research Questions

The challenge offers teams freedom to define the scope and specificity of their research questions and explore development cooperation effectiveness from multiple perspectives.

Participants can approach diverse research questions, ranging from broad macro-level trends to sector-specific investigations looking into areas such as healthcare, gender equity, climate resilience, infrastructure development, among others. The key is to uncover meaningful insights that may not be otherwise captured in mainstream or dominant data sources.

Some examples include:

**Aid Flows & Patterns**

* What factors influence ODA allocation decisions?
* How do economic shocks affect aid commitments?
* What are the emerging trends in sector-specific aid?

**Sentiment & Perception**

* How does media coverage affect public support for aid?
* What is the relationship between domestic politics and aid commitments?
* How does aid sentiment vary across different regions?

**Policy & Impact**

* How do electoral cycles influence aid policies?
* What is the relationship between aid and policy reform?
* How do multilateral vs. bilateral aid patterns differ?

**Effectiveness & Outcomes**

* What factors contribute to aid effectiveness?
* How does aid impact different development indicators?
* What are the spillover effects of targeted aid?

# Challenge Tracks

**Aid Flow Analysis & Prediction**

* Primary Dataset: OECD CRS Database
	+ Detailed project-level data on official development assistance (ODA), with fields including donor and recipient countries, financial amounts, sectors, and aid modalities.
* Example Complementary Data:
	+ World Bank Development Indicators
		- A comprehensive dataset containing socioeconomic indicators such as GDP growth, literacy rates, and poverty levels.
		- Relevance: Provides context on the development needs and economic capacity of recipient countries to predict aid flow patterns.
		- Example Use: Correlate ODA flows with GDP per capita growth in recipient countries.
	+ IMF Economic Data
		- Covers global macroeconomic indicators like inflation, government debt, and trade balances.
		- Relevance: Helps identify economic conditions influencing donor priorities or recipient eligibility.
		- Example Use: Assess how macroeconomic stability impacts sector-specific aid allocation.
	+ UN Comtrade Data
		- Global trade statistics by country, product, and year.
		- Relevance: Highlights trade relationships that may correlate with bilateral aid flows.
		- Example Use: Explore how donor-recipient trade volume aligns with development aid priorities.
	+ IATI Registry
		- Real-time aid activity data from donors, NGOs, and development agencies.
		- Relevance: Complements historical data by offering up-to-date information on ongoing projects.
		- Example Use: Validate predictive models with current aid allocation trends.
	+ World Bank Global Financial Development Database
		- Data on financial inclusion, banking systems, and credit access.
		- Relevance: Examines how financial systems in recipient countries interact with ODA investments.
		- Example Use: Investigate the relationship between ODA for infrastructure and private sector growth.
* Challenge: Develop models to analyze patterns in aid allocation and predict future trends
* Example Projects:
	+ Predictive modeling of aid flows based on economic indicators
	+ Network analysis of donor-recipient relationships
	+ Sector-specific aid distribution patterns

**Development Cooperation Sentiment Analysis**

* Primary Dataset: OECD CRS Database
* Example Complementary Data:
	+ MediaCloud Media Database
		- Focused on specific topics, geographies, and timelines for deep media analysis.
		- Relevance: Allows thematic exploration of media coverage for ODA-related initiatives.
		- Example Use: Identify keywords and topics driving public discourse around ODA.
	+ Social media data
		- Real-time data from platforms like X, Reddit, Instagram, Facebook
		- Relevance: Captures public opinion and sentiment trends regarding development cooperation.
		- Example Use: Analyze public sentiment shifts during major aid announcements or crises.
	+ News archives
		- Historical articles from major news agencies (e.g., Reuters, AP).
		- Relevance: Tracks longitudinal trends in how development cooperation is reported.
		- Example Use: Examine shifts in coverage before and after significant policy changes.
	+ Public opinion surveys
		- Datasets like Pew Research or Gallup polls on development cooperation.
		- Relevance: Captures public attitudes towards aid effectiveness and donor motives.
		- Example Use: Compare public opinion with actual ODA allocation patterns.
* Challenge: Analyze public sentiment and media coverage of development cooperation
* Example Projects:
	+ Media sentiment analysis around major ODA initiatives
	+ Cross-country comparison of development cooperation perception
	+ Topic modeling of development discourse

**Policy & Electoral Impact**

* Primary Dataset: OECD CRS Database
* Example Complementary Data:
	+ Electoral Databases (e.g., IFES Election Guide)
		- Provides data on elections, including dates, results, and voter turnout.
		- Relevance: Links electoral timing with ODA commitments and disbursements.
		- Example Use: Analyze spikes in aid commitments during election years.
	+ Policy Documents
		- Official government white papers and policy statements on ODA.
		- Relevance: Offers insights into policy goals shaping aid distribution.
		- Example Use: Correlate policy changes with shifts in sectoral aid priorities.
	+ Legislative Records
		- National and international legislative debates and bills related to ODA.
		- Relevance: Highlights political discourse and legislative influence on aid decisions.
		- Example Use: Identify legislative drivers of increased or reduced aid to specific regions.
	+ Government Statements
		- Public announcements and speeches by donor country officials.
		- Relevance: Tracks shifts in donor strategies and foreign policy objectives.
		- Example Use: Compare stated priorities with actual ODA flows.
	+ World Bank Political Stability Index
		- Measures political stability and absence of violence/terrorism.
		- Relevance: Assesses how political stability in recipient countries impacts aid allocation.
		- Example Use: Investigate whether stable countries receive more long-term aid commitments.
* Challenge: Examine relationships between ODA, political decisions, and electoral outcomes
* Example Projects:
	+ Analysis of electoral cycles' impact on aid commitments
	+ Policy change effects on aid allocation
	+ Political discourse analysis around development cooperation

**Development Effectiveness Analysis**

* Primary Dataset: OECD CRS Database
* Example Complementary Data:
	+ SDG Progress Reports
		- Reports tracking progress toward Sustainable Development Goals (SDGs) at the national and global levels, including indicators and country-level performance.
		- Relevance: Provides a framework to assess the alignment of ODA with global development priorities and its effectiveness in advancing specific SDGs.
		- Example Use: Analyze how ODA flows align with SDG targets in recipient countries and identify gaps in funding for underachieved goals.
	+ Project evaluation data
		- Results from evaluations of development projects, including outcome metrics, success rates, and qualitative assessments.
		- Relevance: Offers granular insights into the actual impact of ODA-funded initiatives, beyond financial disbursements.
		- Example Use: Investigate trends in successful project outcomes and assess the factors contributing to or hindering their success.
	+ Demographic and Health Surveys
		- Surveys providing detailed data on population health, fertility, mortality, education, or other well-being indicators in developing countries.
		- Relevance: Links ODA spending to improvements in demographic and health-related outcomes.
		- Example Use: Correlate ODA allocations in the health sector with changes in key indicators, for example child mortality rates or access to clean water.
* Challenge: Assess the effectiveness and impact of ODA
* Example Projects:
	+ Impact assessment models
	+ Aid effectiveness predictors
	+ Cross-sector spillover analysis

# Datasets

Participants will be provided with the pre-cleaned OECD Development Assistance Committee Creditor Reporting System (DAC CRS) data for Republic of Korea as a **primary dataset.** Alternatively, participants may also use the unprocessed (raw) OECD DAC CRS data if they wish to. Additionally, participants should use the dataset in conjunction with publicly available datasets. Participants may also explore alternative and unconventional data sources. Some examples include, but are not limited to:

* UNDP’s Human Development Reports: Data on human development indicators.
* New Development Bank Reports: Data on latest projects and annual progress.
* African Development Bank Data: AfDB provides access to information on all its lending projects from 1967 up to now.
* Asian Infrastructure Investment Bank (AIIB) Reports: Information on project status.
* ECLAC (Economic Commission for Latin America and the Caribbean) databases: Statistical databases and publications.
* World Bank Open Data: Datasets on global development, including economic indicators.
* OECD Stat: International economic and social statistics.
* Global Health Observatory (WHO): Public health data.
* Open Data for Resilience Initiative (ODRI): Disaster risk reduction data.
* Earth Observation Data (NASA, ESA): Satellite data for climate and environmental monitoring.
* OpenStreetMap data: Geographic data.
* Kaggle Datasets: A repository of open datasets across multiple domains, including international development.

# Questions and Support

* Development experts: Professionals from USPC to advise on practical use cases.
* Please email info.kr@undp.org in case of any technical questions related to this assignment.

# Submission Requirements

**Submission Documents**

* (1) GitHub repository containing:
	+ Source code
	+ README.md file including:
		- \*Participants may add additional sections and information as they see fit.
		- **Project Overview**
		- **Setup Instructions**
			* \*Code documentation that allows others to understand and reproduce your work
		- **API Documentation**
		- **Analysis Methodology**
			* **\***Explanation of analytical techniques and their appropriateness
		- **Datasets Choice Justification**
			* Primary dataset
			* Secondary dataset
			* \*Explain data sources and processing methods
		- **Key Findings**
			* \* Concise summary of key insights and their significance for development cooperation
			* \*Explanation of how the tool can be used by UNDP
		- **Technical Decisions**
		- **Future Possibilities**
			* **\***Discussion of limitations and potential improvements
* (1) Written proposal specifying a two-month timeline for further prototype development in collaboration with USPC. If the team is selected, this will form the basis for the milestones as stipulated in the agreement to be entered with USPC. (See Annex for proposal template)
* (1) Demo video walkthrough of the initial prototype (max. 5 minutes)

***Submission Process***

1. Create a private GitHub repository
2. Share access with: <https://github.com/sejoonlim-undp>
3. Send email to info.kr@undp.org with your idea proposal attached as PDF and link to demo video to confirm that you have completed the challenge once done.

***Core Requirements***

**Analytical Focus**

1. **Data Integration & Insight Generation**:
	* Meaningful integration of OECD CRS Database with at least one complementary dataset
	* Clear demonstration of statistically significant findings or patterns
	* Identification of non-obvious relationships or trends in development cooperation
2. **Methodological Approach**:
	* Well-documented analytical methodology
	* Appropriate statistical techniques for the chosen research question
	* Transparent handling of data limitations and assumptions

**Technical Implementation**

1. **Data Processing & Analysis**:
	* Effective data cleaning and preparation procedures
	* Appropriate analytical approaches (statistical analysis, machine learning, data visualization, etc.)
	* Robust handling of data quality issues (missing values, outliers, etc.)
2. **Visualization & Communication**:
	* Clear, informative visualizations that effectively communicate key insights
	* Accessible presentation of complex findings
	* Interactive elements that allow for exploration of the data (where appropriate)

**AI Usage**

Encouragement is given for the use of AI tools (GitHub Copilot, ChatGPT, Claude, etc.) to accelerate development. Document which AI tools you used and explain your prompting strategy if relevant.

**Bonus Features (Optional)**

Choose any of these to showcase advanced capabilities:

**Analytical Depth**

* Advanced statistical modeling (regression analysis, time-series forecasting, etc.)
* Machine learning approaches (clustering, classification, anomaly detection, etc.)
* Natural language processing for text data
* Network analysis of aid relationships
* Geographic information system (GIS) integration
* Counterfactual analysis or causal inference

**Technical Innovation**

* API development for data access
* Automated data pipeline for regular updates
* Interactive dashboard for exploration
* Reproducible research notebook
* Novel visualization techniques
* Cloud-based deployment

**Impact & Application**

* Clear policy recommendations based on findings
* Decision support framework for aid allocation
* Monitoring and evaluation framework
* Scenario planning tool
* Comparative analysis across different contexts
* Translation of technical insights into actionable intelligence

**Note on Technical Approach**

We encourage diverse technical approaches that best serve your analytical goals. This hackathon values **analytical insight and relevance over technical complexity**. Your solution may use any programming language, framework, or tool that effectively addresses the challenge. The most important criteria will be the validity, significance, and usefulness of your insights for enhancing development cooperation.

**Guidance for Proposal**

The proposal is fully at the discretion of the participating teams. USPC offers the following guidance on some approaches the teams may consider as they prepare their proposals.

1. **First approach: Advanced Prototype Submission**
Teams submit an advanced prototype that is highly functional. The proposal includes milestones and deliverables that are considered minor changes and refinements.
2. **Second approach: Initial Prototype Submission**
Teams submit an initial prototype that is functional but has considerable room for improvement. The proposal is a comprehensive plan that clearly specifies how the prototype will be substantively improved during the prototype development stage.

USPC offers this guidance, but teams do not need to adhere specifically to these approaches and may formulate their submissions as they see fit. Final selections will be made based on the quality of work and evaluation criteria, regardless of their approach.

# Award Model

The selected teams will win an agreement with USPC to further develop their initial prototype. The agreement money will be given to the team in instalments based on milestones and deliverables described in the proposal submitted by each team. During the two months of the Further Prototype Development stage, teams are not expected to work for USPC full-time. Teams will work independently and may distribute their work hours and deliverables as they see fit, under the understanding that the final product meets the expectations outlined in the agreement. The agreement amount will be inclusive of any applicable taxes and duties that any of the parties may incur.

The teams who are selected to undertake the prototype development process will be required to enter into an agreement with USPC, which will include more detailed conditions pertaining to the prototype development.[[1]](#footnote-2)

# Evaluation Criteria

|  |  |  |
| --- | --- | --- |
| **Evaluation Criteria** | **Details** | **Score Distribution** |
| Proposal Fit | To what extent does the proposed solution address the problem statement effectively? | 30 |
| Data Integration & Methodology | How comprehensively and creatively has the team used the ODA dataset or other relevant data?Did the cross-dataset approach reveal insights beyond what a single dataset could provide? | 30 |
| Analytical Insights | Did the solution uncover novel insights or findings?How accessible did they make complex data and insights? | 20 |
| Potential Impact | Are the findings actionable and implementable? Does the approach show potential for scaling?How does this submission push the boundaries of data analysis in development studies? | 20 |

# Outcomes & Next Steps

* **Collaboration opportunities:** Top teams could be offered further collaboration from development organizations to implement their solutions, such as integration with existing development data platforms.
* **Publication:** Solutions can be featured in development sector reports and publications, opportunities to present at events or conferences.
* **Networking:** Participants gain the opportunity to network with professionals in international development.

**Annex A**

**2025 Data Dive for Development Idea Proposal**

\*Please develop your proposal within three pages. Applicants may use visualizations as they see fit. (Exceeding three pages will result in disqualification)

\*Please delete blue explanations after completion.

|  |  |
| --- | --- |
| **Project Name** |  |
| **Names of Team Members** |  |
| **Description of the Solution** |  \*Brief description of the solution and potential to support UNDP activities. |
| **Further Development Plan & Timeline** | \*Provide a detailed breakdown of development activities on a bi-weekly basis over 8 weeks tailored to your specific project. Examples are provided in blue, but you may change the timeline and activities included as you see fit.**Week 1-2: Analysis and Planning****Week 3-4: Core Functionality Enhancement****Week 5-6: User Experience and Performance****Week 7-8:** Refinement and Launch Preparation |
| **Additional Comments (Optional)** | \*Share any additional information that was not covered in other sections of the proposal |

1. Please refer to <https://popp.undp.org/procedure/innovation-challenges> for more information on UNDP Programme and Operations Policies and Procedures. [↑](#footnote-ref-2)