METHODOLOGICAL NOTE

This document provides an overview of the methodological steps taken to classify impact evaluation (IE) studies within the context of UNDP Signature Solutions. These solutions include Poverty and Inequality, Governance, Resilience, Environment, Energy, and Gender Equality.

To start, we accessed a global repository known as the [3ie Development Evidence Portal](#), which houses over 12,000 IEs and Systematic Reviews. These rigorous pieces of evidence shed light on practices in international development within low- and middle-income countries. It's important to note that the studies in the 3ie repository are categorized according to 3ie's standards, spanning various sectors (e.g., health, education, agriculture, transportation) and themes (e.g., human development, rural development, social development). These categories differ significantly from UNDP Signature Solutions. As a result, it was necessary to reclassify all these papers to align with our corporate thematic areas.

After downloading the available studies, we organized them into a comprehensive [Excel file](#). This file contains information for each study, including the type (IE or Systematic Review), title, abstract, publication year, authors, and publication type (e.g., journal article, academic thesis, study protocol). We then proceeded to clean the dataset by focusing on studies published in the last two decades (from 2003 to 2023). This involved removing potential duplicates, study protocols (as they only outline future Impact Evaluations and lack evidence of outcomes), and ongoing studies. Additionally, we included IEs conducted by UNDP itself, resulting in a total of 12,210 studies.

Next, we embarked on the task of classifying these studies using a data-labeling tool called [Doccano](#). This tool enabled us to label the abstracts of Impact Evaluations according to UNDP Signature Solutions (Poverty and Inequality, Governance, Resilience, Environment, Energy, and Gender Equality), as well as health, which is the most frequent topic in the context of available IEs on international development within low- and middle-income countries.

Initially, a subset of the abstracts was manually classified on Doccano. The team aimed to classify a significant number (at least 150) of studies in each category, adhering to the taxonomy presented in the annex. Based on the labelled dataset, we trained a supervised machine learning algorithm in Python to classify the remaining abstracts. The decision to include categories beyond UNDP Signature Solutions was driven by the recognition that many studies would pertain to these additional categories, thus facilitating subsequent algorithm-based classification.

Tables 1 and 2 present the statistics generated by the machine learning classification. The model demonstrated high precision but lower recall. Precision indicates the frequency with which a machine learning model correctly predicts the target class, while recall measures the model's ability to identify all instances of the target class. Overall, the results are acceptable. Therefore, while the model excels in accuracy, it may sometimes miss identifying instances of the target class.

Table 1. Results of the machine learning classification

Category	precision	recall	fscore	support
Economy	0.81	0.51	0.63	41
Education	0.71	0.79	0.75	19
Energy	0.71	0.75	0.73	16
Environment	0.9	0.82	0.85	115
Gender	0.81	0.49	0.61	43
Governance	0.7	0.4	0.51	65
Health	0.98	0.87	0.92	183
Poverty	0.79	0.69	0.74	67
Resilience	0.83	0.17	0.29	29
Total (Macro)	0.80	0.61	0.67	578
Total (Weighted)	0.86	0.69	0.75	578

Table 2. Precision of the model estimated on a random sample of 109 studies

	# of IE studies
Total	109
Misclassified	6
Governance	2
Gender	1
Environment	0
Energy	0
Resilience	2
Health	2
Poverty	0
% of misclassified/total	5.50%

Annex: Taxonomy used to classify IE studies according to UNDP's signature solutions

Poverty and Inequality
Definition: all policies to support inclusive growth and poverty alleviation over the short to medium-long term.
Keywords: employment, poverty reduction, inequality, social protection, informality, cash transfer, social benefits, increase income.

Environment
Definition: Interventions that contribute to reducing greenhouse gas emissions and/or to reducing vulnerability to climate change through resilience-building and adaptation activities. Any activity/project component that has climate change co-benefits assigned to it.
Keywords: environment, access to water, biodiversity, climate change adaptation, green economy, ecosystem, climate change mitigation, plastics, waste, deforestation, irrigation infrastructure.

Energy
Definition: Energy is a property of objects which can be transferred to other objects or converted into different forms. Energy can be found in a number of different forms. It can be chemical energy, electrical energy, heat (thermal energy), light (radiant energy), mechanical energy, and nuclear energy.
Keywords: energy access, renewable energy, energy transition, clean cooking, electricity access, access to electricity, electricity consumption, grid, energy efficiency, clean energy, energy poverty, energy conservation, reduce energy intensity.

Governance
Definition: Helping countries address emerging complexities by "future-proofing" governance systems through anticipatory approaches and better management of risk.
Keywords: governance capacities, transparency, digital transformation, governance, digitalization, corruption, accountability, risk management, access to justice, rule of law, community involvement.

Resilience
Definition: Supporting countries and communities in building resilience to diverse shocks and crises, including conflict, climate change, disasters and epidemics.
Keywords: resilience, shocks, conflict prevention, conflict, peacebuilding, disaster early warning, migration, recovery, risk management, Covid, community, rural-urban income disparity, microfinance, weather shocks, urban livability/urban resilience.

Gender Equality
Definition: Gender refers to the social, behavioral and cultural attributes, expectations and norms associated with being male or female. Gender equality refers to how these factors determine the way in which women and men relate to each other and to the resulting differences in power between them.
Keywords: women, violence, gender, gap, vulnerable women, women health, gender equality, women's groups, gender differences.

Health
Definition: Health is a state of complete physical, mental and social well-being and not merely the absence of disease or infirmity. Listed below are activities that contribute to improving health, nutrition and population service coverage and outcomes as well as financial protection.
Keywords: health, HIV, Malaria, AIDS, Tuberculosis, health inequalities, Covid