DATA CAPACITY BUILDING IN THE GLOBAL SOUTH

Emergent Patterns and Insights from 24 IDRC Data for Development (D4D) Projects

Strengthening data capacity across civil society, governments, and the private sector in the global south has been an important target outcome of IDRC's Theory of Change for Open Data for Development (OD4D). This study reflects a full review and synthesis of 24 projects related to data capacity building that were undertaken for the purpose of identifying common themes (patterns), effectiveness criteria, and program design considerations. The goal is to determine keys for success, longer-term impact, and expanded knowledge sharing/re-use.

January, 2022

Report

Data Capacity Building in the Global South:
Emergent Patterns and Insights from 24 IDRC Data for Development (D4D) Projects

Executive Summary

Strengthening data capacity across civil society, governments, and the private sector in the global south has been an important component of IDRC's Theory of Change and Programming for (O)D4D. However there is limited evidence on how effective, efficient, and useful the efforts to build such capacity have been in empowering organizations, regulators, educators, and other key actors to be able to fully understand the implications of using data for public good.

This study conducted a full review and synthesis of a selection (~24) IDRC projects related to Data Capacity Building and has identified common themes *(patterns*), effectiveness criteria, and program design considerations that are key for success, longer-term impact and more effective sharing/re-use of knowledge outputs and outcomes.

The Methodology employed the following convergent streams of work:

A. Literature Review on Data Capacity Development and related TOCs to develop a conceptual framework

B. Content analysis and knowledge synthesis from IDRC Project Approval Documents (PAD) and Project Completion Reports (PCR) together with other supporting materials i.e. interim / final technical reports, project outputs, etc.

C. *Internal* consultations (webinar) with the OD4D and D4D network members

The study derived an "Analytic Framework for Capacity Development" that demonstrated utility as an analytic lens for examining the DCB projects.

Key aspects of the Framework were informed by the review of literature on Capacity Development. This included the treatment of capacity as a complex construct that combines individual competencies with organizational capabilities to ultimately deliver capacity as the overall ability of a system to perform. For purposes of sustainability, the framework places a greater emphasis on Capacity Development as a *process*, rather than the more common notion of Capacity Building as a *means*. Furthermore, the explicit articulation of the Unit of Analysis for Capacity Development interventions (i.e. Enabling environment, Sector, Entity, Individuals/Community), emerged as an important program design consideration.

Some of the key insights and best practice recommendations that emerged from the analysis of the 24 projects include:

i. **Capacity Building** for leadership in its broadest sense (*not just political leadership*) is an important factor in the sustainability of Capacity Development

ii. The insertion of external expertise *(through **mentorship or fellowship** models*) into a project, process or institution has proved to be an important catalyst for sustainable capacity development – by facilitating knowledge transfer, two-way learning and while enabling partners on the ground to develop capabilities and define program priorities that matter to them

iii. **Capacity assessment** should be programmed at one or more levels: *Context, Organizational, Individuals*, as a required activity for every Capacity Development project

iv. The deliberate **assessment of context** at the project design and planning stage is critical to effective risk management and mitigation

v. **Institutions create sustainability**: Projects that have established institutions and networks as part of their interventions, contribute significantly to long-term sustainable capacity development

vi. **Policy interventions** are essential for the institutionalizing / sustainability / adaptability of capacity development outcomes

vii. Capacity development interventions should explicitly target one or multiple levels of intervention: ***individual/community***, ***entity***, ***sector/network*** or ***context*** as an explicit design consideration

viii. Categorizing the primary intent of capacity development initiatives as **Means – End - Process** is important for design considerations, and help to inform the appropriate mix of activities and strategies contemplated during project planning

ix. The ultimate goal of any Data Capacity Development program or intervention is to create value and impacts from Data. Explicit consideration of the **Data value chain** helps to embeds data production, use and impact in the overall program design

Aside from its demonstrated utility as an analytic lens for examining the DCB projects, the derived Analytic framework can be incorporated as a guiding tool for capacity development program design and project planning. Potential applications include:

I. A structured questionnaire that covers the dimensions of the framework could be used to guide researcher(s) during the project proposal phase to deliberate on the key inputs, strategies and activities as well as target outputs, outcomes and impact.

II. Integration with the **D4D Resource Centre** currently being established, as a Query interface that would enable researchers to readily locate re-usable resources within any of the elements of the framework e.g. assessment instruments, Open Educational Resource (OER) content, Policy templates, open data, etc.

TABLE OF CONTENTS

1. BACKGROUND AND STUDY RATIONALE 1

2. STUDY METHODOLOGY 2

3. LITERATURE REVIEW ON DATA CAPACITY BUILDING/DEVELOPMENT 3

3.1 Situating Capacity Building/Development 3
3.2 Capacity – Units of Analysis 4
3.3 Capacity Building: Means – End – Process 5
3.4 Capacity Development as Process – *Theory of Change* 7

4. ANALYTIC FRAMEWORK FOR CAPACITY DEVELOPMENT INITIATIVES 8

4.1 Generic Capacity Development Framework 8
4.2 Feedback from Internal stakeholder consultation 10
4.3 Integrating the Data Value Chain 10

5. CONTENT ANALYSIS AND SYNTHESIS FROM IDRC PROJECTS (X24) 12

5.1 Content Analysis 12
5.2 Content Synthesis, Discussion and Insights 15
5.3 Final Project Evaluation – Individual PCR Insights 18

6. DATA CAPACITY DEVELOPMENT – INSIGHTS & RECOMMENDATIONS 23

6.1 Analytic Framework for Capacity Development 23
6.2 Emergent Best Practice 24
6.3 Design Considerations 25
6.4 Limitations of the Research 25

7. BIBLIOGRAPHY 26

ANNEX 1: TABLE 4: SAMPLE OF DATA CAPACITY BUILDING PROJECT LISTING (X 24) 27

ANNEX 2: DESCRIPTIVE META-ANALYSIS OF 24 DATA CAPACITY BUILDING PROJECTS 31

ANNEX 3: INSIGHTS FROM CONSULTATION SESSIONS WITH D4D AND OD4D NETWORKS 34

1. Background and Study Rationale

Data for Development (D4D) has been defined as a key corporate priority for IDRC in 2020 and will seek to leverage the significant body of work accumulated in IDRC programming in order to help achieve the outcomes defined in IDRC's new corporate strategy 2030.

Strengthening data capacity across civil society, governments, and the private sector in the global south has been an important component of IDRC's Theory of Change and Programming for (O)D4D. However there is limited evidence on how effective, efficient, and useful the efforts to build such capacity have been in empowering organizations, regulators, educators, and other key actors to be able to fully understand the implications of using data for public good. Furthermore, the definition of *capacity* must go beyond just the development of technical skills to include everything organizations, regulators, educators, and other key actors need to be able to fully understand the implications of using data for public good.

In order for future programming to contribute to the development and effective application of data in the Global South, IDRC is seeking to leverage the strong foundation of expertise, thought leadership and learning on Data for Development accumulated over the past 10 years across a range of programs to contribute to knowledge mobilization and the sharing of effective data practices with a diverse set of internal and external stakeholders.

This study was designed with the **specific objective** to: Develop design recommendations for consideration with regard to any new IDRC programming to strengthen the capacity of NGOs to use data science to advance initiatives for the public good in the Global South. It will also contribute to a more **general objective** of: Informing IDRC's Strategic Data Capacity Building Research Framework.

This study employs a *Theory of Change* approach to derive a descriptive analytic framework. While the *Theory of Change* is, more often than not, presented in the literature as an approach to program design and evaluation, it is used in this study to provide a consistent and coherent lens through which to conduct the individual project content analysis and synthesize the collective insights. This method enabled the researchers to undertake a systematic content review of a selection (~24) of IDRC projects related to Data Capacity Building in order to identify effectiveness criteria, design elements, critical factors for success, longer-term impact, etc. The grounded insights from this analysis will be used to inform engagements with donor organizations and key stakeholders in the public and private sectors to explore how to scale Data Capacity Building in regional and local non-government organizations (NGOs) working for the public good.

The remainder of this report is structured as follows:

- **Section 2:** Study methodology
- **Section 3:** Literature review on Data Capacity Building/Development
- **Section 4:** Analytic framework for Capacity Development initiatives
- **Section 5:** Insights from consultation sessions with D4D and OD4D networks
- **Section 6:** Content analysis and knowledge synthesis from IDRC projects (x24)
- **Section 7:** Discussion, insights and recommendations

2. Study Methodology

The study was conducted through the following overlapping phases:

A. Developing a conceptual analytic framework derived from a literature review on Capacity Development and a preliminary content analysis of 24 selected IDRC projects
B. Using the derived capacity development analytic framework, conduct a detailed content analysis and knowledge synthesis from IDRC Project Approval Documents (PAD) and related supporting materials (i.e. interim/final technical reports, project outputs, project evaluation reports) to inform Program design considerations and recommendations
C. Incorporating feedback from *internal* stakeholders/experts within the OD4D and D4D Network

This process is represented in the schematic in Figure 1 and then each step is described further.

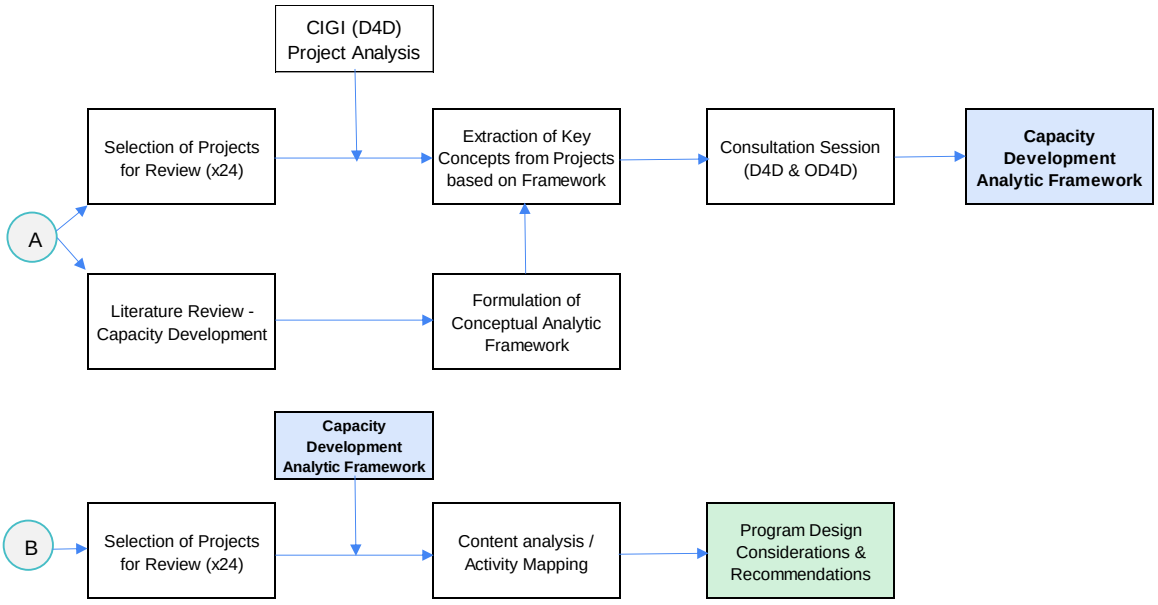

Figure 1: Study Methodology

1. **<u>Selection of Projects</u>** - The Centre for International Governance Innovation (CIGI, 2021) conducted a prior study of 95 IDRC Data for Development (D4D) projects. The meta-analysis included whether each project included a Capacity Building component. The list contained approximately 68 entries tagged with "Capacity Building" as a component. The classification done by CIGI also included information about key words that represent the *Specific Objectives* of the project – these key words included “Data” and “Capacity”. Therefore, further filtering was done on this to reduce the number of projects to 22. Subsequently, two additional projects were added to the sample. The list of the final 24 projects is contained in Appendix I.

2. **<u>Framework for Content Analysis</u> –** To provide a consistent and coherent lens for conducting the content analysis, a conceptual model/framework was developed. This framework was derived from a review of the existing literature, specifically focused on Capacity Building/Development and Data for Development. Using a *Theory of Change* approach the framework was organized into categorized dimensions within the Inputs, Activities, Outputs and Outcomes phases. The framework

also reflected the target of the capacity building intervention (i.e. whether the initiatives were focused at the individual/community level, the entity level, the sector/network level or the context level). This *target* variable identifies the unit of analysis.

3. **Consultation (Internal)** - The initial conceptual model used to conduct the content analysis was presented to and discussed with key stakeholders/experts within the IDRC's D4D and OD4D networks. Based on the feedback from these engagements, the final version of the Capacity Development Framework to be used for the project content analysis was derived.

4. **Content Analysis of Projects** – This step involved analyzing the content of the available project documents for each of the selected 24 IDRC DCB projects using the completed framework as the lens. These documents included project approval documents (PAD), Technical Reports, and in a few cases, Programme Completion Reports (PCR). Each project was mapped to the various dimensions of the analytic framework using a *subjective* scale of 1 to 5, depending on the researchers' assessment of the relative emphasis and extent of the corresponding activities carried out by the project.

5. **Interpretation of the Content Analysis** – The aggregate mapping of the 24 projects and the emergent patterns were analyzed to draw inferences about program design and primary activities, as well as about any apparent relationship with other project attributes. These grounded findings, insights, and inferences are discussed in Section 7.

3. Literature Review on Data Capacity Building/Development

3.1 Situating Capacity Building/Development

Capacity building (or capacity development) has been an *evolving* and sometimes *elusive* concept in the international development discourse and has often been used as an umbrella term, referring to a wide spectrum of activities ranging from training and workshops to previously separate development approaches - such as "*organizational development*", "*community development*", "*integrated rural development*", and "*sustainable development*" (Gillespie, 2005).

The terms, capacity building and capacity development, have been used interchangeably in literature, but capacity development seems to have become the preferred term in more recent work. According to Horton's (2002) reasoning, the word '*development'* seems to better fit the intended meaning, suggesting a more organic process of growth and development than does '*building'*, which connotes an externally planned or engineered approach.

It is perhaps useful, at this juncture, to define three related terms (**Capacity, Capabilities, Competencies)**, sometimes used interchangeably within the context of capacity development, therefore clarity of meaning will be important for our subsequent analysis.

An ECDPM Study on Capacity, Change and Performance (Zinke, 2006) offers the following definitions linking the three (3) constructs:

- **Capacity** - the overall ability of a system to perform and sustain itself: the coherent combination of *competencies* and *capabilities*.

- **Capabilities** - a broad range of collective skills of organizations or systems which can be both 'hard' (e.g. policy analysis, marine resource assessment, financial resources management) and 'soft' (e.g. the ability to earn legitimacy, to adapt, to create meaning and identity). Capabilities can be understood as the building blocks of an organization's overall *capacity* to perform.
- **Competencies** - the skills and abilities of individuals often characterized as *knowledge*, *skills* and *attitudes* that combine with other resources, assets and processes to determine organizational capabilities.

With this hierarchy of constructs in mind, we adopt the definition of *Capacity Development* offered by Baser and Morgan (2008):

> *"that emergent combination of individual competencies, collective capabilities, assets and relationships that enables a human system to create value"*

3.2 Capacity – Units of Analysis

Capacity is a complex construct that exists not only in individuals but also between them in the institutions and social networks they create (Fukuda-Parr et al., 2002). As a result, capacity development planning and interventions have to be deliberate in the unit of analysis and may target multiple levels of intervention, such as 'individuals', 'groups', 'organisations', or 'societies'.

Bolger (2000) expresses a similar principle as being, an important attribute of capacity development, and defines the following units of analysis:

- **The Enabling Environment**: represents the broad context within which development processes take place. Attempts to effect change at the enabling environment level generally take a considerable length of time given the nature of the issues being addressed - *policies, structures, attitudes, values etc*. While not all capacity development initiatives will seek to effect change in the enabling environment, they will need to be sensitive to factors at this level.

- **The Sector/Network Level**: developing countries and donor agencies are increasingly focusing their investments on this level, e.g. sector or sub-sector programs. This reflects an increasing awareness of the importance of coherent sector policies, strategies and programming frameworks, as well as effective coordination within and across sectors. Reforms at this level can contribute significantly to synergies and promote more effective use of existing capacities.

- **The Organizational Level**: this capacity level focuses on organizational structures, processes, resources and management issues. An important dynamic exists among the organizational, the sectoral and enabling environment levels. Similarly, organizational performance depends on the availability, effective use and motivations of individuals.

- **The Individual Level**: this level in the Capacity Development framework refers to individuals as social or organizational actors (e.g. small holder farmers, water engineers, planners, accountants, etc.) and their knowledge, skills, and attitudes. From a Capacity Development perspective, change at the individual level should be contemplated as part of a broader setting. Too often, development projects have focused narrowly on training of individuals without giving adequate attention to

organizational issues, broader processes of empowerment, or relevant factors in the 'enabling' environment.

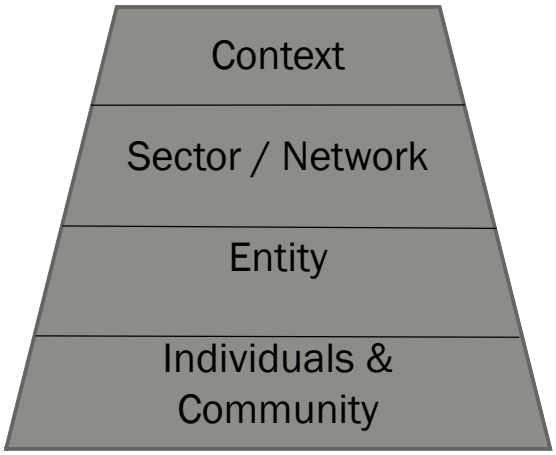

Figure 2: Capacity Development - Units of Analysis

3.3 Capacity Building: Means – End – Process

Acknowledging the complexity of Capacity Development, the international donor community has become increasingly aware that capacity is not something that can delivered from the outside, but is instead something that must be "*wilfully acquired*" over time (Lavergne & Saxby, 2001). New modalities of Capacity Development reflect a shift away from *hard* capacities like equipment and infrastructure (*Means*) and an increased interest in *softer* capacity enablers like leadership, strategy, and support for good governance (*Process*).

Program design must, therefore, contemplate Capacity Building/Development's intent as one or a combination of the following:

- **Capacity Building** as *Means*
  - *Partners abilities are strengthened to carry out specific activities within the project/program, ensuring that the requisite skills and resources are in place to support the interventions*
- **Capacity Building** as *End*
  - *Core abilities of partner organisations are strengthened to enable greater local ownership of problems and solutions, and to fulfil the organisation's own goals and mission, as defined by the organisation and the local context*
- **Capacity Development** as *Process*
  - *Emphasis on adaptability and sustainability. Development of relationships and networks for mutual benefit. Creation of on-going knowledge exchange to support partner's viability and lead to on-going adaptation, learning, and change*

Table 1 illustrates examples of projects categorized according to their design intent: Means-End-Process. The listing of 24 projects contained in Appendix I reflect the complete categorization. The summary shows that the projects analyzed are relatively evenly distributed across the 3 categories: End (9), Means (7) and Process (8). However there is a stronger correlation between the design intent and the levels of intervention. Capacity Development projects with a "Process" intent are frequently designed at the "Context" level while those targeting entity or sector level are often designed as "Means".

The program design logic here is clear: when designing Capacity Development programs for long-term sustainability, one should explicitly target the context. Otherwise, entity or sector interventions, are often designed to deliver the "requisite skills and resources" as a Means towards achieving explicit sector or entity objectives.

Table 1 - Examples of Project Classifications

MEANS	END	PROCESS
Using Data for Improving Education Equity and Inclusion (109372)	Strengthening District Health Management Team capacity to use health information systems data and to engage stakeholders to address teenage pregnancy (108936)	Strengthening the use of open data in francophone Africa to improve policy, citizen engagement and drive innovation (109525)
Data Use Innovations for Education Management Information Systems in The Gambia, Uganda, and Togo (109371)	Scaling open data for development in Latin America (108490)	Strengthening School Leadership Towards Improving School Resiliency (109563)
Mobile technology and enhanced counselling to improve family planning among Syrian refugees and host communities in Lebanon and Jordan (36 months) (109089)	Digital new deal for Africa (109358)	Making a Feminist Internet: the Feminist Internet Research Network (108598)
Strengthening National Health Information Systems in the Middle East towards Evidence-Informed Decision Making in Health Systems (108917)	Preparing Haitian Youth for Digital Jobs (108360)	Building an Africa Open Data Network (108492)

3.4 Capacity Development as Process – *Theory of Change*

***Capacity Development**: that emergent combination of individual competencies, collective capabilities, assets and that enables a human system to create value*

Theory of Change is an explicit expression of program logic that seeks *to* articulate the mechanism of change linking the Inputs and Activities to short-term Output/Outcomes and long-term Goal/Impact. The previously adopted description of Capacity Development as a *Process* emphasizes sustainability and mechanisms that lead to on-going adaptation, learning, and change. This articulation lends itself readily to conceptualization, design, and evaluation using the *Theory of Change (TOC)* concept.

While the *TOC* is typically used as an approach to program design and evaluation, for the purposes of this study, we employ the approach to derive a descriptive analytic framework that can be used to provide a consistent and coherent lens through which to conduct the individual project content analysis and synthesize the collective insights. As Birckmayer et al (2000) reflect, a Theory of Change does not have to be *right* to be *useful*.

Figure 3: Theory of Change - Components

The review of the literature on Capacity Development, as well as the initial review of the 24 project documents, identified the most commonly appearing elements (*concepts*) associated with each stage of the TOC (see figure 4). These are used to define the categorized dimensions of our analytic framework in the following section.

IMPACT	Achieving Development Objectives
OUTCOMES	Change -> *values, attitude, skills, norms, knowledge, behaviour* Effectiveness -> *development oriented targets* Capabilities -> *act and commit, deliver, relationships, coherence* Adaptability -> *the ability to adapt and self-renew* Sustainability -> *not dependent on external partners for ongoing support*
OUTPUTS	Competencies, Transfer/acquisition of skills/knowledge, Policies, Infrastructure, Governance frameworks, Systems
ACTIVITIES / PROCESS	Training, Education, HRD, Advocacy, Policy development, System development, R&D, Resource Mobilization
INPUTS	Assessment, Awareness, Strategies, Resources, Policies, Lessons learned, Best practice, Data, External expertise

Figure 4: Capacity Development TOC Elements

4. Analytic Framework for Capacity Development Initiatives

4.1 Generic Capacity Development Framework

By consolidating the TOC elements from figure 4 into a tractable set of dimensions, the generic analytic framework for Capacity Development was derived as shown in figure 5. A definition of each dimensional construct is provided in the subsequent table.

This framework can be used in several ways:

- Comparative evaluative lens for Capacity Development projects (see content analysis of projects in Section 6)
- Guiding framework for Capacity Development program design (See program design recommendations – Section 7)

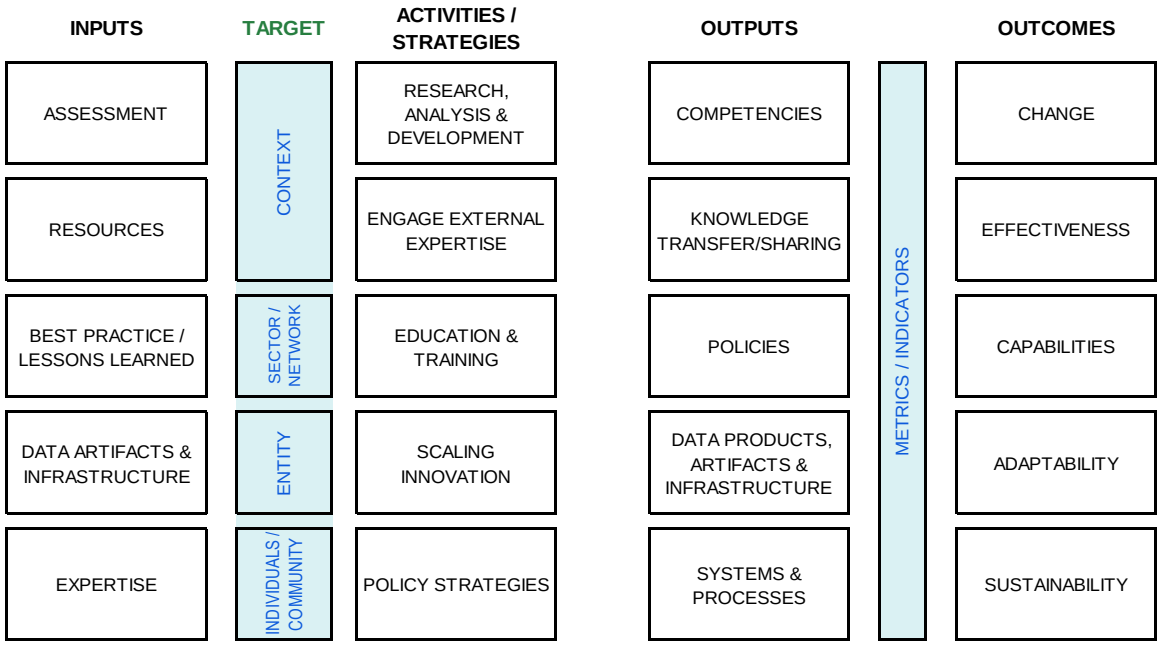

Figure 5: Analytic Framework for Capacity Development

Table 2: Data Capacity Development Framework - Glossary of Terms

Construct/Term	Definition
INPUTS	Tangible: Financial, Human and Material Resources as well as Intangibles e.g. Information, Data, Knowledge, etc.
Assessment	A structured approach for analyzing capacity across three dimensions: *individuals, organizations and the enabling environment.*
Resources	Financial, Human and Material resources.
Best Practice / Lessons Learned	Using experiences of previous projects as input into the existing project.
Data Artifacts & Infrastructure	Data and/or derivatives as a resource input: may involve surveys or retrieval of administrative data or the use of remote sensing methods
Expertise	3rd party expert resource e.g. consultant that provides input services to the project.

ACTIVITIES / STRATEGIES	Set of approaches and actions employed in the project to transform inputs into outputs
Research, Analysis & Development	Investigative activities or study regarding a particular problem; Also includes analyzing data to extract useful information, or development of data products such as models
Engage External Expertise	Engage 3rd party expert who coaches local persons for a limited period of time. Institutional twinning, coaching and mentoring programmes are typical activities.
Education & Training	Training is a planned intervention to modify attitude, knowledge or skill behavior through a learning experience, while Education is more concerned with the general growth and development of beneficiaries.
Scaling Innovation	Implement a relevant strategy during the project or immediately after the project's end that leads to widespread use of an innovation… growing the innovation's impact to match the level of need.
Policy Strategies	Activities that engage with and encourage policy actors e.g. Capacity Building initiatives with government stakeholders or partnerships with international organizations or governments to promote policy impact.
OUTPUTS	What will be delivered from completing activities? What will stakeholders "*Expect to See*"
Competencies	The knowledge, skills and attitudes of individuals.
Knowledge Transfer / Sharing	The deliberate and planned exchange of knowledge as well as to ad hoc sharing of knowledge
Policies	A law, regulation, procedure, administrative action, incentive, or voluntary practice of governments and other institutions.
Data Products, Artifacts & Infrastructure	Data products are products built through the use or analysis of data/An application or tool that uses data to help entities improve their decisions and processes/A product that facilitates an end goal through the use of data.
Systems & Processes	A process is a series of steps or actions performed to achieve a particular goal while systems are what's used to execute the process
OUTCOMES	What others will do with the outputs. What do stakeholders "*Want to See*".
Change	Changes in values, attitude, skills, norms, knowledge, behavior of the target beneficiaries.
Effectiveness	Ability to realize development oriented targets.
Capabilities	A broad range of collective skills of organisations or systems which can be both 'hard' and 'soft'. Capabilities can be understood as the building blocks of an organisation's overall capacity to perform.
Adaptability	Ability to adjust and self-renew in response to novel and/or complex emergent situations.
Sustainability	Not persistently depending on external partners for support.

TARGET / METRICS / INDICATORS

These elements are not stages per se, but are included in the Analytic Framework (Fig 5) to highlight the importance of their explicit consideration in Capacity Development planning and interventions. As discussed previously in Section 3.2, such initiatives should, at inception, be deliberate in specifying the unit of analysis and how they target one or more levels of intervention, such as 'individuals', 'groups', 'organisations', or 'societies'. Likewise, each stage and element (activities, outputs, outcomes) should have standardized metrics that enable more effective tracking of data Capacity Development *(qualitative and quantitative)* outputs and outcomes.

4.2 Feedback from Internal stakeholder consultation

Internal stakeholder consultation sessions were conducted with members of the IDRC-convened Data for Development (D4D) and Open Data for Development (OD4D) Networks, in the format of two webinars conducted on March 23rd and 25th, 2021. The sessions were used to present the preliminary insights and findings from the study, and in particular, the proposed Capacity Development Analytic Framework. Objectives of these sessions were to gain:

- Critical feedback on the utility of the conceptual framework for Data Capacity Development.
- Comment on lessons learned and preliminary program design parameters.

A summary of the discussions and commentary from these sessions is presented in **Annex 2.** Some of the key feedback and suggestions include:

- General consensus that the "means", "ends" and "process" categorization was a useful consideration in the Capacity Development component of the projects; it would be interesting to analyze further the level of complexity of different Capacity Development initiatives where some have core capacity objectives and others only a capacity component as a *means* to some other primary objective.
- The data value chain is a well-established Theory of Change concept that should be incorporated into the framework. We have integrated the stages/concepts from the data value chain into the analytic framework, in order to enhance its utility.
- Explicit consideration of metrics/indicators should be incorporated into the framework. Each stage and element (activities, outputs, outcomes) should have standardized metrics that enable more effective tracking of data capacity development (qualitative and quantitative outputs and outcomes). A portfolio of metrics/indicators and assessment tools aligned with standardized measurement frameworks such as GDB and the SDGs should be made available to program planners and implementers.

4.3 Integrating the Data Value Chain

The data value chain (Open Data Watch, 2018) describes the process of data production and use from first identifying a need for data to its final use and possible reuse. The data value chain (DVC) has four major stages: *collection, publication, uptake, and impact* (figure 6*)*. These four stages are further separated into twelve steps: *identify, collect, process, analyze, release, disseminate, connect, incentivize, influence, use, change, and reuse*.

The DVC shows the complex set of steps from data creation to use and impact and can be used as a tool to monitor and evaluate the data production process. In this regard, it provides a form of *Theory of Change* logic relating to Data Production and Use/Impact. Although our generic analytic framework was conceptualized for Capacity Development, the ultimate goal of any Data Capacity Development program or intervention is to create value and impacts from Data. Hence there is an opportunity to integrate concepts from the data value chain into the analytic framework, in order to enhance it's utility, particularly when applied to D4D projects contemplating data capacity development as a *Means* or *Process*.

We map the elements of the DVC to our generic analytic framework as illustrated in fig 6 and described in the Table 2 below:

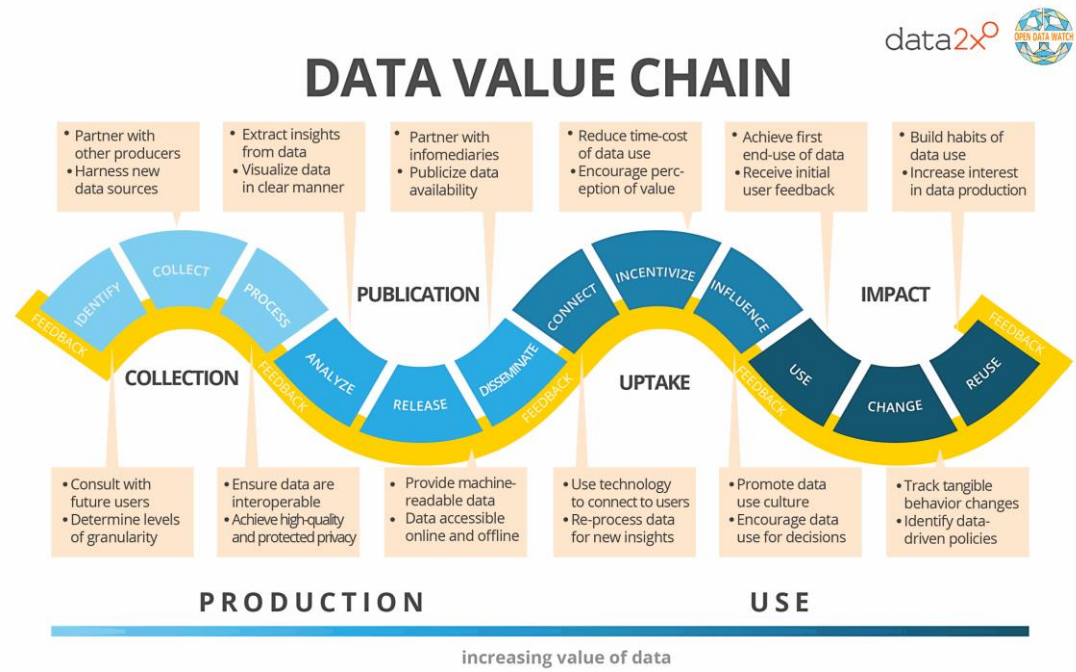

Figure 6: Data Value Chain - Prepared for Data2X by Open Data Watch

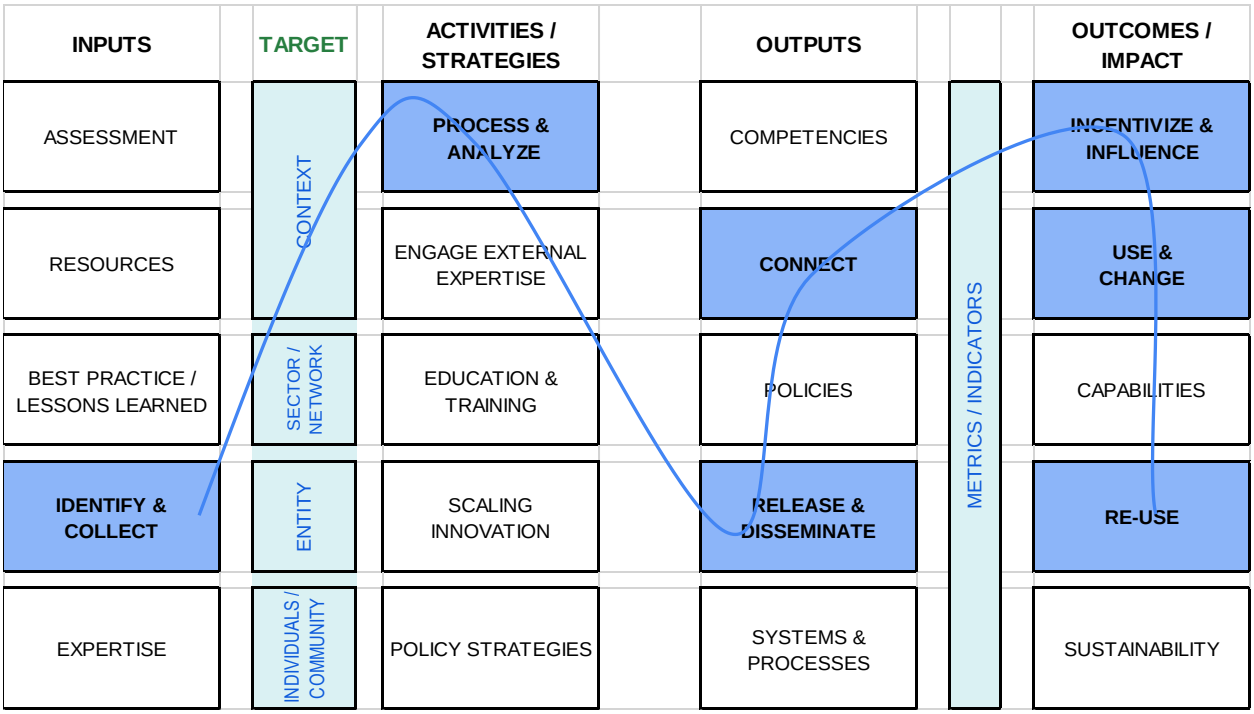

Figure 7: Data Capacity Development (DCD) Analytic Framework

Table 3: Mapping of DVC elements to Analytic Framework

Stage	DVC Element		Framework Element	Rationale for Mapping
Collection	identify	identifying what data to collect and how we will use them	DATA ARTIFACTS & INFRASTRUCTURE	The identification and collection of Data and/or derivatives can be a key resource input for capacity buuilding initiatives
	collect	may involve surveys or retrieval of administrative data or the use of remote sensing methods		
	process	ensure data is correctly recorded, classified, and stored in formats that allow further use	RESEARCH, ANALYSIS & DEVELOPMENT	Research or analysis of data to extract useful information, or development of data products such as models
Publication	analyze	analyzing data to extract useful information		
	release	publish data and the accompanying metadata with appropriate documentation in online and offline formats	DATA PRODUCTS, ARTIFACTS & INFRASTRUCTURE	These data products, artifacts and supporting infrastructure can be typical outputs of capacity building initiatives
	disseminate	deliver data to intended users through appropriate dissemination channels;		
Uptake	connect	use various channels to connect data to users eg. trainings, seminars or by improving the user experience offered by websites, data portals, etc.	KNOWLEDGE TRANSFER/SHARING	Connecting data to users through various knowledge sharing/transfer mechanisms
	incentivize	encourage perception of value; reduce transaction cost of using data	CHANGE	Changes in behavior of the target beneficiaries is a primary objective of DCB initiatives; incentivizing or promoting aaaaaaa data use culture are important means for doing so
	influence	promote data use culture; encourage users to incorporate data into the decision-making process		
Impact	use	achieve first end-use of data; facilitate data use to understand a problem or make a decision	EFFECTIVENESS	Effectiveness in our framework relates to the ability to realize development oriented targets; Effective use and tangible behaviour change are typical types of outcomes
	change	impact of data: change the outcome of a project or improving a situation; tangible behaviour change or data-driven policies		
	reuse	build habits of data use; combining them with other data and sharing them freely	ADAPTABILITY	Effective re-use of Data will enhance its adaptability and re-purposing for novel and/or complex emergent situations

The elements highlighted in blue reflect the corresponding steps/activities derived from the Data Value Chain. This conceptual integration illustrates that while the DVC is a subset of the Data Capacity Development Analytic framework, it provides a useful connecting tissue throughout, embedding data production, use and impact in the overall DCD Theory of Change.

5. Content analysis and synthesis from IDRC Projects (x24)

This section describes the use of the derived analytic framework to conduct the content analysis of 24 IDRC capacity building projects.

5.1 Content Analysis

Annex 2 provides a descriptive meta-analysis of the 24 projects selected for the study. This includes attributes such as: Region, Grant size, target sector, IDRC Programming, etc. Beyond this descriptive summary, the content analysis carried out using the derived analytic framework results in a much more nuanced examination of the Theory of Change logic for each project, as well as of the intent implied in the general and specific objectives. We also examine outputs and outcomes from both technical reports and project completion reports *(where available)*.

The analysis proceeded as follows:

a. The content of each project was mapped to the various dimensions of the analytic framework, using a *subjective* scale of 1 to 5, depending on the researchers' assessment of the relative emphasis and extent of the corresponding activities carried out by the project.
b. The mapped projects were then visualized individually and in aggregate, as well as scored to determine any distinctive emergent patterns associated with different groupings of projects.

To demonstrate the nature of the output of this analysis, the following three projects were used as an illustration (figures 8a-8c) with the scored circles representing the assessed weight of emphasis in each of the dimensions of the analytic framework. Each of the example projects represents one of the types of Capacity Development: End-Means-Process:

- **Preparing Haitian Youth for Digital Jobs (108360) - *End***
- **Using Data for Improving Education Equity and Inclusion (109372) - *Means***
- **Strengthening School Leadership Towards Improving School Resiliency (109563) - *Process***

In the first case, *108360 - Preparing Haitian Youth for Digital Jobs*, the primary objective was to build digital skills capacity among young women in Haiti (*END*). The target of the intervention is accordingly reflected primarily at the level of *individuals/community*, but also *context* as the project sought to improve the degree of Internet connectivity and the level of technical network skills in Haiti.

The focus of the primary Capacity Building efforts were on *training*, and the designed curriculum and courses were informed by considerable *research* into the prevalence of online job types. The primary outputs of this Capacity Building project were the improved digital competencies of the 300 young women trained, which it is hoped will lead to changes in values, attitudes, skills and ultimately behavior changes in the community. There were also efforts towards sustainability in the creation of a local enterprise and related systems and processes that could facilitate access to online jobs by the trained young women. The mapping in figure 8a illustrates this project narrative.

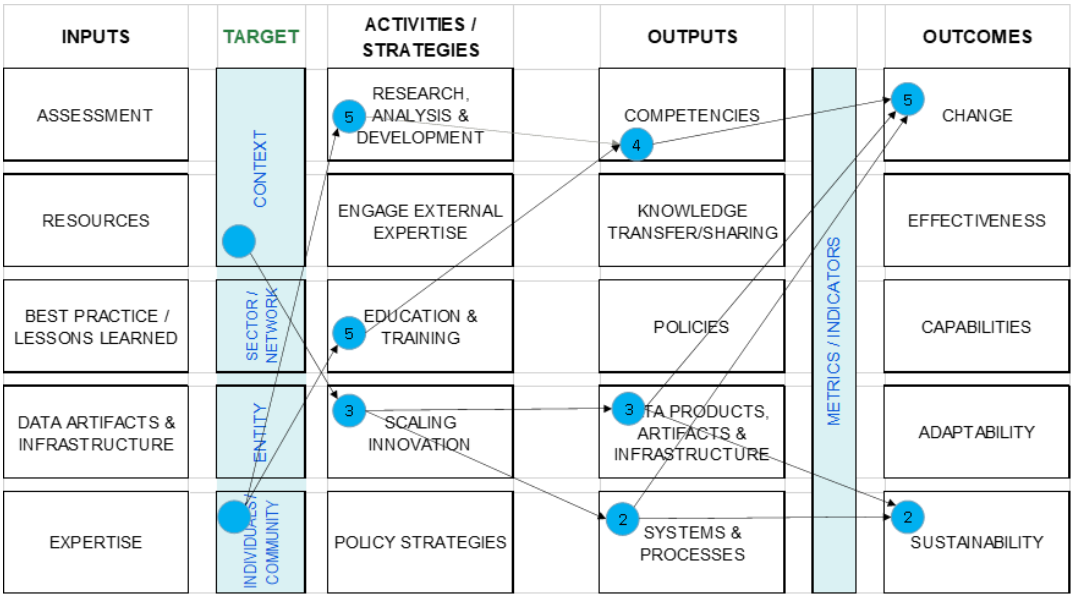

Figure 8a: Mapping - Preparing Haitian Youth for Digital Jobs (108360)

In the second case, *109372 - Using Data for Improving Education Equity and Inclusion*, the Capacity Building purpose was to develop data skills in order to support the primary objective of using data by government agencies for education sector analysis and policy. The target of the intervention is therefore reflected primarily at the level of organizations (*entities*) and the *education sector*. The primary Capacity Building activities focused on *training* in order to develop and ultimately deliver the capabilities within

ministries of education and other education stakeholders to effectively analyze education sector data, and influence *behavior change* about equity and inclusion in the education sector. There was also an emphasis on scaling the initiative through *data-driven policies.* The mapping in figure 8b illustrates this project narrative.

NB: The elements highlighted in blue are labelled using the corresponding steps/activities derived from the Data Value Chain, compared to the more generic labelling used previously. Both representations yield the same analysis results.

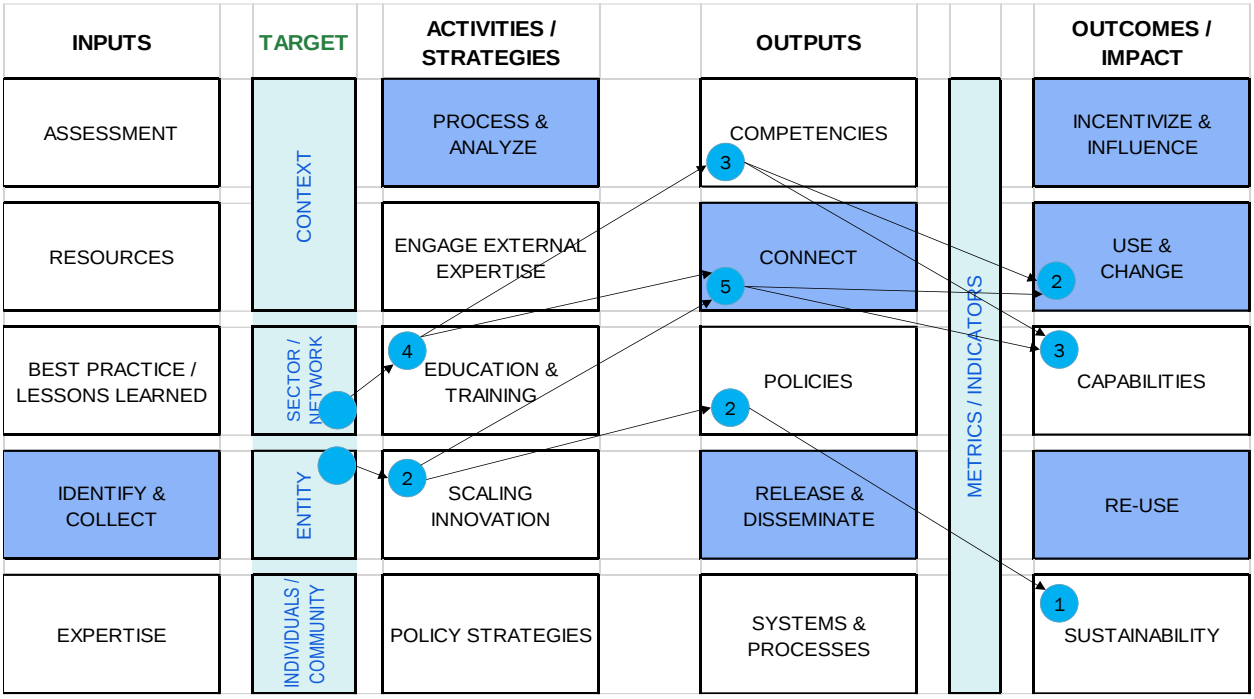

Figure 8b: Using Data for Improving Education Equity and Inclusion (109372)

In the third case, *108598 - Making a Feminist Internet: the Feminist Internet Research Network*, illustrated in figure 8c, the Capacity Development purpose was to support learning and build capacity within a research network to generate evidence and inform policy and practice relating to the needs of women and gender diverse people in internet policy processes and in decision-making overall. The target of the intervention is therefore reflected primarily at the level of community and the wider country context. The primary Capacity Building activities focused on *education*, *training* and *research, and data analysis* to inform *policy* and practice. Primary outputs are research competencies, and the release and dissemination of information that can influence policies, decision-making and action, relating to women's online rights and feminist frameworks. The focus on building the capacity of a network puts the outcome emphasis on *knowledge-sharing*, *adaptability* and longer-term *sustainability.* The mapping in figure 8c illustrates this project narrative.

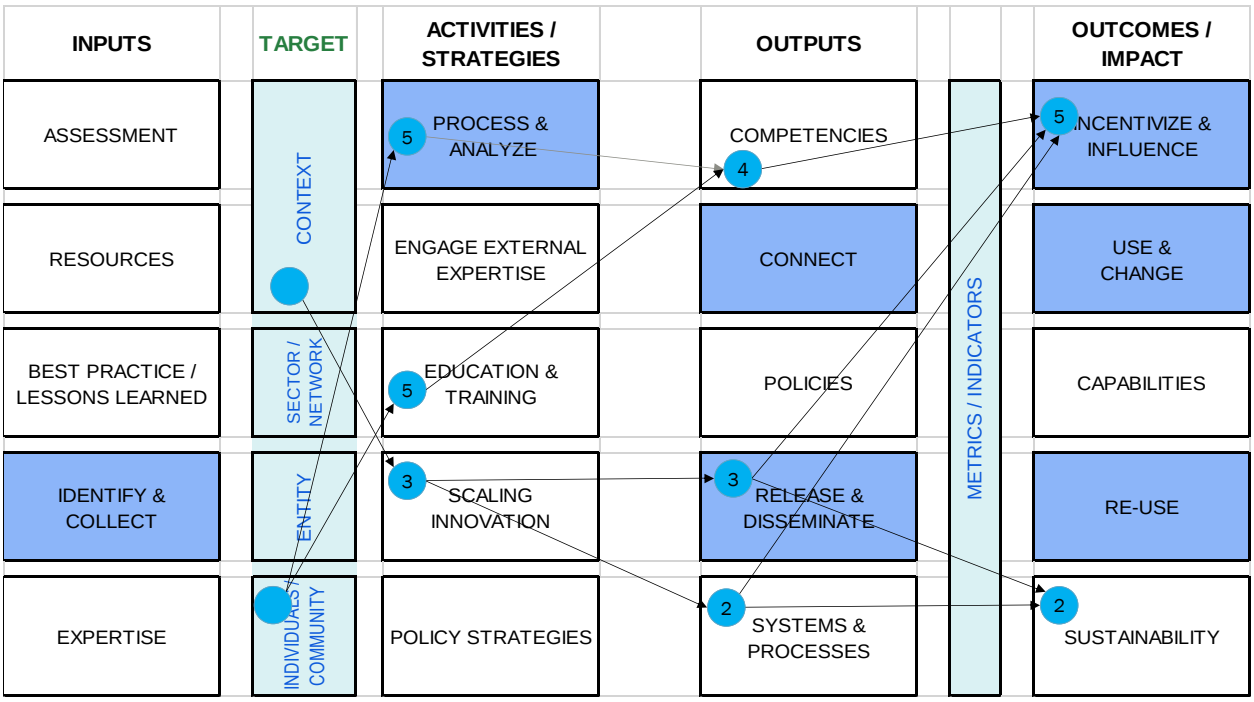

Figure 8c: Making a Feminist Internet: the Feminist Internet Research Network (108598):

All three selected examples are very different types of Capacity Development interventions. The visual mapping of the chain of actions from intent and activities through outputs and outcomes illustrate diverse emergent patterns and demonstrates the utility of the analytic framework in analyzing and representing the different approaches.

5.2 Content Synthesis, Discussion and Insights

This analysis and mapping was conducted for all 24 projects and the synthesis is visually represented in figure 9a. The number and size of the circles reflect the degree of emphasis in the synthesis of the content for the 24 projects. Additionally, figure 9b represents a synthesis of the results of mapping in terms of the aggregate weighted emphasis across the various dimensions within each of the four stages of the framework: Inputs – Activities – Outputs – Outcomes.

INPUTS	ACTIVITIES / STRATEGIES	OUTPUTS	OUTCOMES
ASSESSMENT	PROCESS & ANALYZE	COMPETENCIES	INCENTIVIZE & INFLUENCE
RESOURCES	ENGAGE EXTERNAL EXPERTISE	CONNECT	USE & CHANGE
BEST PRACTICE / LESSONS LEARNED	EDUCATION & TRAINING	POLICIES	CAPABILITIES
IDENTIFY & COLLECT	SCALING INNOVATION	RELEASE & DISSEMINATE	RE-USE
EXPERTISE	POLICY STRATEGIES	SYSTEMS & PROCESSES	SUSTAINABILITY

Figure 9a: Framework Mapping of 24 IDRC Capacity Development projects

INPUTS	ACTIVITIES / STRATEGIES	OUTPUTS	OUTCOMES
ASSESSMENT 3%	PROCESS & ANALYZE 36%	COMPETENCIES 22%	INCENTIVIZE & INFLUENCE 22%
RESOURCES 8%	ENGAGE EXTERNAL EXPERTISE 4%	CONNECT 32%	USE & CHANGE 13%
BEST PRACTICE / LESSONS LEARNED 50%	EDUCATION & TRAINING 39%	POLICIES 17%	CAPABILITIES 33%
IDENTIFY & COLLECT 25%	SCALING INNOVATION 13%	RELEASE & DISSEMINATE 24%	RE-USE 8%
EXPERTISE 14%	POLICY STRATEGIES 7%	SYSTEMS & PROCESSES 5%	SUSTAINABILITY 24%

Figure 9b: Framework Analysis of 24 IDRC Capacity Development projects

The following observations are worth noting:

Inputs: The highest weighting was for *"Best Practice/Lessons Learned"* which was not surprising given the design of many of the projects as well as the way the approval template is structured. Many of the projects under consideration referenced previously funded projects and these are stated in the proposal. In terms of the template for approval of projects (PAD), it includes a section "Application of Lessons Learned" where the relevant previously funded projects are identified. This is important both for continuity and also being able to leverage knowledge resources accumulated from prior projects. It underscores the importance of current efforts to establish a searchable, indexed Knowledge Resource Centre where knowledge artifacts can be readily located. Another common/frequently occurring input was *Data Artifacts and Infrastructure (Identify & Collect)* – a number of the projects have adopted/adapted existing tools and technologies. For example project 108678 is focused on the adoption of an existing Health Information System, while project 109354 refines a previously proposed Global Pluralism Index.

A visible area of limited focus in terms of input was "*Assessment*". Capacity assessment is not a consistent feature of project inputs and the literature suggests that it should be. We believe capacity assessment should be programmed at one or more levels: *Context, Organizational, Individuals*, as a required activity for every Capacity Development project and should be identified as a part of the proposal submission. Existing measurement frameworks such as the Global Data Barometer could become important references for this input activity, especially as it relates to Capacity-Building relating to Data Governance (See Section 7.3). It has already been noted in the literature that the success of Capacity Development initiatives is dependent on context and therefore it is important to understand the environment within which the study/project is to be executed.

Activities/Strategies: The two most commonly occurring activities across the projects were (i) *Education and Training* and (ii) *Process & Analyze (Research, Analysis and Development)*, as might be expected for Capacity Development projects. "*Process & Analyze"* activities ranged from data management and analysis initiatives to product development and research studies.

One of the important points emerging from the consultation with the stakeholders (D4D and OD4D) is the gap in the effective sharing/reuse of the outputs produced from these projects. For example, several projects that involved significant education and training activities included a curriculum development component, but there were no mechanisms in place to make these training artifacts readily shared/ available for re-use.

Therefore although "*Education & Training*" and "*Process & Analyze*" were prevalent project activities/strategies, there could be more sharing of the learning and knowledge artifacts (i.e. outputs) produced as a result of these activities. The correct "*Systems and Processes"* need to be established to facilitate this sharing (e.g. a "resource centre" that is currently being contemplated).

Outputs: Outputs across the project analyzed appeared to be relatively evenly distributed across *Competencies*; *Connect (Knowledge Transfer/Sharing)*; and *Release & Dissemination (Data Products, Artifacts & Infrastructure)*;

Policies and Systems *& Processes* appeared to have less emphasis but are important mechanisms for institutionalization and sustainability of these initiatives.

The outputs from these projects are important, tangible artifacts that should be available to share and re-use by other researchers and data*4*development initiatives in the Global South. This shared re-use represents an important opportunity for sustained Capacity Development in this region and mechanisms for doing so should be an integral component of project design.

A number of the projects have as an output the creation of networks focused on the Global South. This becomes an important aspect of Capacity Development as a process, whereby these networks can/should continue after the life of the project and foster ongoing Knowledge Transfer and Sharing that will in turn promote Capacity Development.

Outcomes: Of the intended outcomes emphasized by the 24 projects, development of *Capabilities* was most prevalent, with *Sustainability* also expressed as an important goal. The ability to *Incentivize and Influence* the behavior of target Actors was also a prevalent outcome that will help to promote a data use culture. *Adaptability (Re-Use)* was the area least emphasized across the projects, underscoring a need for a deliberate focus on *Data Re-Use* in the project design.

5.3 Final Project Evaluation – Individual PCR Insights

Much of the content analysis and synthesis for this study was based on the **Project Approval Documents (PAD)** which reflect program design intent rather than what was actually accomplished. Where available, this was augmented by a review of other project artifacts such as technical reports and other deliverables.

Project Completion Reports (PCR) were available for four (4) of the projects studied. PCRs provide fairly detailed reviews and independent perspectives on the projects and provide insights into the results and management of the projects as well as highlighting lessons learned. In this section, we reflect on some of the key questions / insights arising from a review of those PCR reports, and relate these findings to the insights derived from the conducted analysis reported in the previous section.

Box 1: Program Completion Reports

Project Completion Reports (PCRs) are the result of reflective exercises that provide detailed reviews and independent perspectives on the projects and provide insights into the results and management of the projects as well as highlighting lessons learned. We are not attempting here to re-evaluate these projects, rather we're pulling out some of the key insights/commentary from the PCRs, and then reflecting on how it correlates with the insights from our own analytic framework.

A. 108490 SCALING OPEN DATA FOR DEVELOPMENT IN LATIN AMERICA

i. Were the risk mitigation strategies outlined in the Project Approval Document effective? Were there any unanticipated risks that affected the project?

The quantity and variety of themes was a risk to producing concrete results in any one of them. Initially it was assumed that partnerships and a thorough knowledge of the data ecosystem in the region would help mitigate this risk. Yet, as the project moved to implementation, it was decided that it would be better to concentrate only on two of the subjects.

This underscores the case for greater emphasis and deliberation on "Assessment" during project design/approval. From the earlier framework analysis, Capacity Assessment is not a consistent feature of project inputs and the literature suggests that it should be. We believe Capacity Assessment should be programmed at one or more levels: *Context, Organizational, Individuals*, as a required activity for every Capacity Development Project. This will also help considerably in Risk assessment and mitigation strategies or more practical project scoping.

ii. Leadership was important. Some members of the project team were later used in other projects.

Leadership in its broadest sense (*not just political leadership*) is an important factor in the sustainability of Capacity Development. A key output of this project has been ILDA's leadership capacity:

- ILDA has been called on to provide technical support for other initiatives.
- ILDA has been key in the development of policies for the region, using their innovative participatory approach that involved key stakeholders.
- ILDA has been able to become a standalone not-for-profit entity following this project.

iii. **The PCR indicates a strong POLICY focus and several important policy documents were produced from this project.**

This policy focus was not apparent in the original project design (PAD), where the focus was more on training and curriculum design. This may have resulted in a change in focus of the project during implementation. However it does underscore the potential usefulness of the analytic framework as a reference during project design/approval to ensure a more holistic, integrated strategy formulation.

B. 107075 FROM DATA TO DEVELOPMENT: EXPLORING THE EMERGING IMPACT OF OPEN GOVERNMENT DATA IN DEVELOPING COUNTRIES

This project provides an excellent example of "**Capacity Development as Process (pp.5)**" and highlights some important elements and outcomes:

i. Effective mentorship brings an opportunity for two-way learning between mentors and mentees, and supports capacity building whilst leaving partners on the ground in the driving-seat, defining research priorities that matter to them

As pointed out by the evaluation, the project has not only focused on funding research, but has also provided mentoring and networking opportunities. Mentors, paired with each research partner, have provided guidance on social science research, content advice and technical knowledge, which has significantly improved the quality of the network's research products.

This particular network structure has allowed partners to share their experiences with Open Data in their national context and gain an understanding of what is happening in different parts of the world - effectively gaining a global perspective on Open Data. This mentorship model might be considered a pre-cursor to The School of Data fellowship programme.

ii. Leadership: One of the main contributions of the project was around building leaders

In the mid-term evaluation, 60% of research partners surveyed have already started working on other open data projects, with other international and local governments and international development funders.

iii. Policy Impact

The project also contributed to large scale positive change through contributions to formulation of Open Data policies across the world. In particular, this research informed directly the instruments currently being used to evaluate Open Government data programs around the world i.e. Open Data Barometer.

C. 108492 - BUILDING AN AFRICA OPEN DATA NETWORK

This project provides another excellent example of "**Capacity Development as Process**" where the emphasis is on the d*evelopment of relationships and networks for long-term adaptability and sustainability.*

i. Fellowship Program

The embedded fellowship model was employed, seeking to build the organisational capacity of civil society organisations, but primarily targeting journalists and the media industry. A key project output that should provide a re-usable educational resource was: A curriculum for data capacity building in gender NGOs.

ii. Assessment of Context

There were substantial delays in obtaining the country clearance from the Kenyan Government, which meant that the LDRI activities (network coordination) started almost one year later. From the earlier framework analysis, we highlighted that Capacity Assessment should be programmed at one or more levels: *Context, Organizational, Individuals*, as a required activity for every Capacity Development project. This will also help considerably in risk assessment and mitigation strategies or more practical project scoping.

iii. Institutionalization for Sustainability

This project helped establish the foundations to enable the use of Open Data for Development (OD4D) in Africa, by building/creating African institutions and networking them. It had more success in West Africa than Southern and East Africa. Establishing the foundations is only the first step along the way to governments adopting and implementing Open Data policies that will benefit citizens.
Some tangible outcomes include:

- Key networks were established among African policy-makers, researchers and civil society
- Led to two (2) OD4D hubs in Africa (Francophone and Anglophone)
- New co-funding partnerships with the Hewlett foundation

D. 108360 PREPARING HAITIAN YOUTH FOR DIGITAL JOBS

This project evaluation identified mixed results. It was not possible to achieve some of the expected results, such as the employment target that was originally intended for the graduates of the program and the employment opportunities and engagement of diaspora did not emerge as was expected.

Although identified as a shortcoming by the researchers, the project leads stated that this was an unrealistic expectation and the real objective of developing courses and curriculum and positively impacting young women in Haiti through digital Capacity Building had been accomplished. This project was therefore categorized as "**Capacity Development as End".**

i. Assessment of Context

This project evaluation insight reiterates the importance of Capacity Assessment, to be programmed at one or more levels: *Context, Organizational, Individuals*, as a required activity for every Capacity Development project as a means of risk assessment and identifying mitigation strategies.

ii. Capacity Development – Multiple Units of Analysis

We reflected (in 3.2) that Capacity Development planning and interventions have to be deliberate in the unit of analysis, but may target multiple levels of intervention, such as 'individuals', 'groups', 'organisations', or 'societies'.

For this project, the focus of the primary capacity-building efforts were on *training individuals and groups,* and the primary outputs / outcomes were the courses and training materials developed, and improved digital competencies of the 300 young women trained. The project was evaluated as successful in this objective. There was also an organizational Capacity Development objective relating to the creation of a local enterprise and related systems and processes that could facilitate access to online jobs by the trained young women. This objective was not achieved. However these multi-level objectives highlight the importance of an integrated planning framework that assists researchers and planners in holistic program design at the outset.

iii. Effective Sharing/Reuse of the Data and Data Products

An important outcome measure of Capacity Development initiatives is the extent of data / knowledge sharing re-use. In this case, the curriculum, courses and training delivery mechanisms developed for the program in Haiti have been adapted and re-used for a larger regional initiative – Caribbean School of Data – which has attracted follow-on funding from Google Foundation.

6. Data Capacity Development – Insights & Recommendations

6.1 Analytic Framework for Capacity Development

Too often, Capacity Building projects have focused narrowly on training of individuals without giving adequate attention to organizational issues, broader processes of empowerment or relevant factors in the 'enabling' environment. We have derived and used throughout this study, an Analytic Framework that contemplates the broader construct of Capacity Development. The framework was used primarily to provide a consistent and coherent lens through which to conduct the individual project content analysis and synthesize the collective insights.

The aggregate mapping of the 24 projects and analysis of the emergent patterns yielded the following insights and recommendations about program design and primary activities:

- "**Capacity Assessment**" has not been a consistent feature of project inputs. The literature suggests that it should be – the success of capacity development initiatives is dependent on context and therefore it is important to understand the environment within which the study/project is to be executed. We recommend that Capacity Assessment be programmed at one or more levels: *Context, Organizational, Individuals*, as a required activity for every Capacity Development project. Existing measurement frameworks such as the Global Data Barometer could become important references for this input activity, especially as it relates to initiatives that seek to enhance Data Governance Capacities.

- Categorizing the primary intent of capacity development initiatives as **Means - End - Process** is important for design considerations. Both the aggregate framework analysis as well as the examination of individual sample projects in each of these categories (see 6.2) reflect different emergent patterns of the Theory of Change logic. This suggests that categorizing projects as *Means - End - Process* can help to inform and guide project designers in terms of the appropriate mix of activities and strategies contemplated during planning. For instance, the application of the data value chain is particularly relevant when considering D4D projects where Data Capacity Development is primarily intended as *Means* or *Process*.

- Explicit consideration of metrics/indicators that would be used to evaluate the impact of the project was not an observed feature of the project approval documents. This was a recurring concern throughout the consultations and has been visually incorporated into the analytical framework at the Output/Outcome/Impact stages, as a prompt that this should be a program design consideration. A portfolio of metrics/indicators and assessment tools could be aligned with the framework and made available to program planners and implementers as part of the D4D Knowledge Resource Center currently being developed.

- The target of the Capacity Development intervention i.e. whether the initiatives were focused at the **individual/community level, the entity level, the sector/network level or the context level**, is an important design consideration. As highlighted in the consultation sessions, the level of complexity of different Capacity Development initiatives varies significantly across projects and it is often difficult to translate individual training interventions into improved capabilities and institutional capacities. A given project may have more than one target level, therefore multiple levels of

intervention at the sector, context or entity should be considered to facilitate sustainability. Categorizing projects in terms of the target of the capacity development intervention can help to inform and guide project designers in terms of the appropriate mix of activities, strategies and related indicators contemplated during planning.

- Explicitly incorporating the well-established **Data Value Chain (DVC)** into the Capacity Development Analytic Framework enhances its utility, since the ultimate goal of any Data Capacity Development program or intervention is to create value and impact from data. It also illustrates (figure 7) that for the sustainability of Capacity Development, other elements beyond the DVC need to be considered.

Aside from its demonstrated utility as an analytic lens for examining the DCB projects, we believe there is merit in using this Analytic Framework for Capacity Development as a guiding tool for capacity development program design and project planning.

A structured questionnaire that covers the dimensions of the framework could be used to guide researcher(s) during the project proposal phase to deliberate on the key inputs, strategies and activities to the target outputs, outcomes and impact. This structured process will help align planned activities with the outputs and outcomes and ensure that sufficient details are specified in the proposal document and in so doing strengthen the project proposal and planning phases. Another potential use-case of the framework could be integration with the **D4D Resource Centre** currently being established, as a Query interface that would enable researchers to readily locate re-usable resources within any of the elements of the framework e.g. assessment instruments, OER content, policy templates, open data, etc.

6.2 Emergent Best Practice

Several recurrent activities have emerged from the project analysis that could be considered as “best practice” for the sustainability of Capacity Development programs.

- **Leadership Capacity building**: Leadership in its broadest sense (not just political leadership) is an important factor in the sustainability of Capacity Development. The emergence of leaders that provide continuity and sustainability of D4D programming has been a feature of successful projects.

- **Mentorship / Fellowship**: The insertion of external expertise into a project, process or institution has proved to be an important catalyst and opportunity for knowledge transfer, two-way learning and while enabling partners on the ground to develop capabilities and define program priorities that matter to them. These are important features of sustainable Capacity Development.

- **Assessment of Context / Risk Management**: The deliberate assessment of context at the project design and planning stage is critical to effective risk management and mitigation. Several project experiences and outcomes underscore the importance and need for proper assessment of the context and conditions, particularly for large-scale initiatives in fragile contexts. This will help to inform more effective risk assessment and mitigation strategies or more practical project scoping.

- **Institutions create sustainability**: Projects that have established institutions and networks as part of their interventions, contribute significantly to long-term sustainable Capacity Development in global south contexts. ILDA is a prime example of an emergent institution that has continued to have sustainable Capacity Development and policy impact in Latin America.

- **Policy Strategies/Outputs**: Policy interventions are essential for the institutionalizing/sustainability / adaptability of Capacity Development outcomes. This was a feature of successful programs, however they were not a sufficiently consistent activity across all projects analyzed. The provision of policy guidelines, templates and strategic options will be important re-usable artifacts in the D4D Resource Centre.

6.3 Design Considerations

A summary of the key program design considerations arising from these insights are as follows:

x. **Capacity Building for leadership** in its broadest sense (*not just political leadership*) is an important factor in the sustainability of Capacity Development.

xi. The insertion of external expertise *(through **mentorship or fellowship** models*) into a project, process or institution has proved to be an important catalyst for sustainable Capacity Development – by facilitating knowledge transfer, two-way learning and while enabling partners on the ground to develop capabilities and define program priorities that matter to them.

xii. **Capacity assessment** should be programmed at one or more levels: *Context, Organizational, Individuals*, as a required activity for every Capacity Development project.

xiii. The deliberate **Assessment of Context** at the project design and planning stage is critical to effective risk management and mitigation.

xiv. **Institutions create sustainability**: Projects that have established institutions and networks as part of their interventions, contribute significantly to long-term sustainable Capacity Development.

xv. **Policy interventions** are essential for the institutionalizing / sustainability / adaptability of Capacity Development outcomes.

xvi. Capacity Development interventions should explicitly target one or multiple levels of intervention: ***individual/community***, ***entity***, ***sector/network*** or ***context*** as an explicit design consideration.

xvii. Categorizing the primary intent of Capacity Development initiatives as **Means – End - Process** is important for design considerations, and helps to inform the appropriate mix of activities and strategies contemplated during project planning.

xviii. The ultimate goal of any Data Capacity Development program or intervention is to create value and impacts from Data. Explicit consideration of the **Data Value Chain** helps to embeds data production, use and impact in the overall program design.

6.4 Limitations of the Research

It should be noted that a number of the projects included in the analysis are still in progress. The project documentation that was available for synthesis included mainly the project proposals (PADs), therefore this analysis was primarily focused on the projects' design intent, as emphasized in each of the project approval documents, together with interim project outputs. While only a few projects had documentation to reflect a post completion review, we had the opportunity to review the PCR reports for four (4) projects and reflect on additional emerging insights.

7. Bibliography

Baser, H., & Morgan, P. (2008). *Study on Capacity, Change and Performance*.

Birckmayer, J. D., & Weiss, C. H. (2000). Theory-based evaluation in practice: what do we learn? *Evaluation Review*, *24*(4), 407–431.

Bolger, J. (2000). Capacity development: why, what and how. *Capacity Development Occasional Series*, *1*(1), 1–8.

Centre for International Governance Innovation (CIGI). (2021). *Data for Development (D4D): Reviewing and Charting the Role of IDRC in the Global Data Agenda - Phase 1 Report*.

Fukuda-Parr, S., Lopes, C., & Malik, K. (2002). Institutional innovations for capacity development. *Capacity for Development: New Solutions to Old Problems*, 1–21.

Gillespie, B. (2005). *Theories of change: exploring of IDRC understandings about capacity development; first draft report*.

Horton, D. (2002). *Planning, implementing, and evaluating capacity development*. Citeseer.

Lavergne, R., & Saxby, J. (2001). Capacity development: Vision and implications. *Capacity Development Occasional Series*, *3*, 1–11.

Open Data Watch. (2018). *The Data Value Chain:Moving from Production to Impact*. https://opendatawatch.com/publications/the-data-value-chain-moving-from-production-to-impact/

Zinke, J. (2006). ECDPM Study on Capacity, Change and Performance-Final Workshop. *Workshop Report, July*.

Annex 1: Table 4: Sample of Data Capacity Building Project Listing (x 24)

Project#	Project Title	General Objective	Region	Target Sector	Unit of Analysis	Capacity Intent
107075	From Data to Development: exploring the emerging impact of open government data in developing countries	The research will address key issues and gaps in knowledge e.g. well-intentioned open data initiatives may cause adverse effects by exacerbating social or economic inequalities. They can favour well-resourced groups that are more capable of extracting value from the data for their own economic and political gain. Commercial elites in Bangalore, for example, have used the digitization of land records to find gaps in title and errors in documentation to gain ownership of land occupied by poor communities. The project will also develop data collection instruments to help explain if and how open data is bringing change to developing countries. It will engage policymakers at global and local levels to foster robust evidence-based practice in this emerging policy area.	International - HQ _ North America	Open Data	Context	Process
108334	Improving Prospects for Data Enabled Livelihoods Among Marginalized Communities	To foster better livelihood opportunities in the digital economy for women and youth in the Middle East and North Africa communities	MENA	Education	Context; Individual	End
108360	Preparing Haitian Youth for Digital Jobs	Create the enabling conditions for young Haitians to find employment in the digital economy in Haiti through addressing skills and infrastructure deficits.	Caribbean	Education	Context; Individual	End
108490	Scaling open data for development in Latin America	To scale open data initiatives in Latin America that contribute to reducing corruption, improving health service delivery, increasing resilience of cities and reducing violence against women.	LATAM	Open Data	Context	End
108492	Building an Africa Open Data Network	To scale the development impact of open data initiatives in Africa, through promoting the adoption of improved open data principles, best practices, policies, partnerships and use.	Africa	Open Data	Context	Process
108598	Making a Feminist Internet: the Feminist Internet Research Network	To gather evidence and inform policy and practice that will ensure the needs of women and gender diverse people are taken into account in internet policy processes and in decision-making overall.	Africa	General	Context	Process

108602	Harnessing Big Data to meet the Sustainable Development Goals: Building Capacity in the Global South	To support research on big data for development that supports developing countries to measure progress on and contributes to the achievement of the Sustainable Development Goals (SDGs).	LATAM;Asia; Africa	General	Context; Individual	End; Process
108678	Reducing Child Mortality: The Role of Mobile Electronic Health Information System	To improve survival of children under five years of age by building and using a mobile phone-based health information system for collecting data directly from both caregivers of children and health facilities, while strengthening the capacity of caregivers to identify and seek services for life-threatening diseases.	Africa	Health	Sector	End
108868	OD4D Phase II	To catalyse the effective release and use of open data in generating greater accountability and transparency in governments, promoting local innovation, and improving delivery of key public services such as education and health.	multiple/international	Open Data	Context	End
108872	Improving Dietary and Health Data for Decision-Making in Agriculture and Nutrition Actions in Africa	To develop efficient solutions to measure indicators that track nutrition and health outcomes and their progress at community and national levels.	Africa	Health/Agriculture	Context; Community	End
108917	Strengthening National Health Information Systems in the Middle East towards Evidence-Informed Decision Making in Health Systems	The overall aim of this project is to promote data-informed decision-making and accelerate achievement of SDGs in Lebanon and Jordan by strengthening the current health information systems and ensuring the necessary processes, tools and competences are in place for the generation, analysis and utilization of relevant, reliable and timely data and information to effectively inform health policies and programs, including those related maternal, child and adolescent health.	MENA	Health	Sector	Means
108936	Strengthening District Health Management Team capacity to use health information systems data and to engage stakeholders to address teenage pregnancy	Develop and strengthen the capacity of district health management teams in conducting research using the District Health Information Management System platform and engaging with key stakeholders in identifying gaps and finding solutions to addressing the problem of teenage pregnancy in the Volta Region of Ghana.	Africa	Health	Entity; Community	End

108998	Intersectional analyses of gender and health of female-headed households in low- and middle-income countries	Promote gender equality by raising the visibility of female-headed households in low and middle income countries (LMICs), quantifying their role in women's and children's health, exploring intersectionalities between gender and other dimensions of inequality, and training researchers and policymakers in gender equity analyses.	International	Health	Individual; Community	Means
109011	Strengthening the Adolescent Reproductive Health Information System in Palestine	Inform current and future Adolescent Reproductive Health- Health Information System (ARH HIS) planning and implementation in Palestine by generating evidence about its technical social and operational drivers, and recommendations to address gaps in data, indicators and their analysis.	MENA	Health	Sector	Means
109089	Mobile technology and enhanced counselling to improve family planning among Syrian refugees and host communities in Lebanon and Jordan (36 months)	To develop, implement, and evaluate contextualized, culturally sensitive, data-driven, and evidence-based strategies to encourage the adoption of quality family planning services among vulnerable host communities and Syrian refugees living in Lebanon and Jordan.	MENA	Health	Sector	Means
109204	AgMIP Adaptation Teams Start-up - A CLARE Transition Activity	To use AgMIP tools and results to inform decision making and planning in the agricultural sectors at the national level in three countries: Senegal, Ghana and Zimbabwe.	North America	Agriculture	Sector	Means
109250	Protecting Personal Data in the Digital Economy in Latin America	Deepen understanding and strengthen capacity by developing and applying citizen data audit tools to study citizen perceptions of how platformization and private use of personal data is impacting the delivery of services to vulnerable communities in five countries in Latin America.	Latam	General	Community	End
109354	Leaving No One Behind: Addressing Inequalities and Exclusions Through the Global Pluralism Index	Over a three-year period, this project will enable the Global Centre for Pluralism to develop the partnerships, capacities and structures to scale the Index to a wide geographic coverage, capturing different kinds of pluralism experiences around the world and engage stakeholders to use the findings to advance more inclusive policies and practices.	International	General	Entity	Process

109358	Digital new deal for Africa	Strengthen research and policy capacity to enable policy leaders in Africa to respond to rapidly changing digital environments and ensure policy ecosystems associated with artificial intelligence, data and the future of work are informed by objective, high quality research.	Africa	General	Context	End
109371	Data Use Innovations for Education Management Information Systems in The Gambia, Uganda, and Togo	To strengthen the capacity of ministries of education in Uganda, Togo and the Gambia to adapt open source health information systems to local and national education sector needs for better policy formulation, planning and implementation.	North America	Education	Sector	Means
109372	Using Data for Improving Education Equity and Inclusion	The general objective is to enhance the use of equity and inclusion data by government for education sector analysis and policy.	International	Education	Entity	Means
109518	Strengthening Inclusive Open Data Systems in Africa and South-East Asia	To better understand the relationships between the availability of inclusive public data, and its use through innovative technologies and collaborative stakeholder engagement to address complex policy challenges and policy change to inform African and Asian policy makers and the broader communities of civil society, the private sector, and local communities and women	Africa/Asia	Open Data	Context	Process
109525	Strengthening the use of open data in francophone Africa to improve policy, citizen engagement and drive innovation	Achieving more inclusive governance and public service delivery in countries in Francophone Africa through strengthening the environment for open and shared data to be used for sustainable development. CAFDO will support multidisciplinary applied research, capacity building and innovations to support the release and use of data by governments and intermediaries such as technologists, journalists, academics, civil society groups and rights organizations.	Africa	Open Data	Context	Process
109563	Strengthening School Leadership Towards Improving School Resiliency	Contribute to improved quality and equity of the continued learning and well-being of girls and boys in the Global South during the prolonged school closures of the COVID-19 crisis, and future emergencies.	International	Education	Sector	Process

Annex 2: Descriptive Meta-Analysis of 24 Data Capacity Building Projects

Referencing the prior study of IDRC Data for Development (D4D) projects. The meta-analysis was conducted on the 24 projects based on the following variables:

Grant Size: This variable is described in the PAD document and states the amount in Canadian dollars funded by IDRC. The coded values of the variables are "Up to 350k", "350-700k", "700k-1mil.", "1-1.5mil.", and "> 1.5mil."

Institute Region: This variable is coded based on the "Institute" variable listed in the PAD. The values are based on the main geographic regions relevant to the sample, namely "Africa", "N America", "S. America", "Europe", "Mena region", "Asia" and "Other/International."

IDRC Programming Areas: This variable is coded based on the overall project content, as described in the PAD, and as related to the "Theme" variable. The values refer to the following IDRC program areas: "Equity in Global Health", "Climate Resilient Food Systems", "Democratic and Inclusive Governance", "Education and Science", "New Prosperity" and "other". Note, a project can fall into more than one value.

Gender: This variable is coded based on the PAD information, when available, or based on the PAD overall content. It follows the IDRC expertise in defining how project investment can address gender issues. The variable values are "gender aware", "gender sensitive", "gender responsive", "gender transformative" and "none."

Strategy: This variable is coded based on the Theory of Change, strategies and milestones described in the PAD and describes key implementation strategies of projects, related to education (e.g., research), engagement of external actors (events), innovation (development and deployment of technology and tools), capacity building (live courses), communication (online portals) and other.

Policy: This variable describes the level of planned engagement with government policy processes. The variable values are "direct engagement with policy maker" (e.g., in-person participation in government committees), "direct participation in policy processes" (e.g., contributing to public consultations), and "indirect objective to influence policy and regulation" (e.g., mentions policies and regulation, but only an indirect plan on how to interact with it).

In addition to these variables derived from the CIGI study, we also analyzed and coded the projects based on the following additional variables:

Unit of Analysis/Intervention: This variable is coded based on the targeted level of the Capacity Building/Development intervention (*as described in section 3.2*) based on the PAD overall content. The variable values are: Context (i.e. the enabling environment), Sector (Sectors, Networks, Communities of Practice), Entity (Organizational level), Individual, Community (general citizen awareness, participation).

Capacity Intent: This variable represented the researchers' interpretation of the Capacity Building/Development's intent as one or a combination of the following: Means-End-Process, as described in section 3.3.

Target Sector: This variable indicated whether the Capacity Development initiatives were focused at the sector level and targeted a specific sector such as Agriculture, Health or Education.

Summary Meta- Analysis

The summary descriptive analysis of the 24 Capacity Building projects contained in the study sample is presented below. Refer to the Project listing in Appendix 1 for the individual details.

Institute Region: The distribution of the institutes supporting the projects in the study sample is shown. The majority are based in Africa.

REGION - Institute	N	%
Africa	9	38%
N. America	3	13%
S. America	3	13%
Mena	4	17%
Asia	2	8%
Caribbean	2	8%
Other/International	7	29%

Grant Size: the majority of projects receive Grant support within the ranges 350k to 1mil with 58% of cases in this category.

GRANT SIZE	N	%
Up to 350k	1	4%
350-700k	7	29%
700k-1mi	7	29%
1-1.5mi	3	13%
.> 1.5mi	5	21%

Target Sector: The majority of the studies (54%) were sector-focused targeting sectors such as Health, Education and Agriculture. Others were more broad-based in their focus, targeting the Policy context or general Capacity Building in topics such as Open Data or Big Data.

TARGET SECTOR	N	%
General	5	21%
Education	5	21%
Health	6	25%
Open Data	6	25%
Agriculture	1	4%
Health/Agriculture	1	4%

IDRC Programming: The distribution of the project sample across the IDRC Program areas is shown. As might be expected, the majority (88%) are associated with Education and Science. A significant number (77%) are also associated Democratic and Inclusive Governance, an area of emphasis for Data for Development.

IDRC PROGRAMING AREAS	N	%
Equity in Global Health	9	38%
Climate Resilient Food Systems	3	13%
Democratic and Inclusive Governance	17	71%
Education and Science	21	88%

Project Strategies: For this select sample, Capacity Building is obviously the primary activity. However, in general projects utilize a combination of approaches from the available portfolio of strategies as the distribution shows. Our subsequent content analysis will unpack this further to identify the dominant strategies employed.

PROJECT STRATEGIES	N	%
Research	23	96%
Engagement of external actors	23	96%
Deployment of technology/tools	12	50%
Capacity Building	24	100%
Scalling innovation/data strategy	22	92%
Communication	17	71%

Policy Strategies: nearly all the projects (92%) reflect some kind of policy intent. Far less (13%) participate directly with policy processes. However most (58%) have some form of direct engagement with policy makers

POLICY STRATEGIES	N	%
Direct engagement with policy maker	14	58%
Direct participation in policy processes	3	13%
Direct objective to influence policy and regulation	22	92%

Unit of Analysis/Intervention: 50% of the projects focus their intervention on Context (the enabling environment). Several projects exhibit multiple levels of intervention. For example all projects in the sample focused on the individual also include context interventions. A good example of this is "Preparing Haitian Youth for Digital Jobs (108360)" which addresses both individual competencies as well as the context in terms of Internet availability.

INTERVENTION	N	%
Context	12	50%
Sector	7	29%
Entity	3	13%
Individual	4	17%
Community	4	17%

Capacity Intent: For a significant number of the projects (42%), Capacity Building is the primary purpose (*End*) of the intervention. eg. Preparing Haitian Youth for Digital Jobs (108360); For 29% of the projects, Capacity Building is the *Means* towards supporting some other project objective. eg. "Using Data for Improving Education Equity and Inclusion (109372)". For 33% of the projects, Capacity Development is more of an organic *Process* eg. "Making a Feminist Internet: the Feminist Internet Research Network (108598)"

APPROACH	N	%
Means	7	29%
End	10	42%
Process	8	33%

Annex 3: Insights from Consultation sessions with D4D and OD4D networks

Internal stakeholder consultation sessions were conducted with members of the IDRC-convened Data for Development (D4D) and Open Data for Development (OD4D) Networks, in the format of two webinars conducted on March 23rd and 25th, 2021. The sessions were used to present the preliminary insights and findings from the study, and in particular, the proposed Capacity Development Analytic Framework. Objectives of these sessions were to gain:

- Critical feedback on the utility of the conceptual framework for Data Capacity Development
- Comment on lessons learned and preliminary program design parameters

Table 5: Feedback & Responses - D4D Consultation Session - 23-03-2021

D4D Meeting – March 23rd		
Topic	**Comments / Feedback**	**Response / Implications**
Means- End-Process	There was general consensus that the "means", "ends" and "process" approach was an interesting and useful consideration in the capacity development component of the projects; It would be interesting to delve into the level of complexity of different Capacity Development initiatives as it varies significantly across projects. Especially as some have core capacity objectives and others only a capacity component.	All 24 projects have been categorized accordingly. The subsequent framework analysis will seek to determine whether this attribute influences the project design considerations. This could be a question / assessment to be included at the outset as the research proposal is being framed
Capacity Building *vs* Capacity Development	Many participants agreed that Capacity Development is more suitable than Capacity Building and should be the preferred term moving forward.	Capacity Building suggests building something new from the ground up, according to a pre-imposed design, while Capacity Development is believed to better express an approach that builds on existing skills and knowledge, driving a dynamic and flexible process of change, borne by local actors[1]. Capacity Development is more applicable to these types of research projects so should be the adopted term.
Knowledge Sharing	Several participants expressed a concern that one of the barriers to learnings from projects is the lack of "sharing" of findings, outputs, experiences and best practices. A number of the projects have as outputs data artefacts that can be shared. However, it is difficult to share if peers do not know what exists. This ultimately leads to data silos – one of the roadblocks for data use identified in the value chain literature.	This is an important consideration that must be addressed. There are parallel initiatives that are addressing this concern e.g. through the development of a resource centre to "house"/"direct" end users to these resources. The aim being to reuse these data products.
Data Value Chain	When looking at Capacity Development in national statistics, people are mostly interested in looking at results. The Data Value Chain, analysing the four major stages: *collection, publication, uptake and impact*, is a well-	The ultimate goal of any Data Capacity Development program or intervention is to create value and impacts from Data. Hence the Data Value Chain which is a Data *Production-Use-Impact* construct becomes very relevant.

1 https://www.europarl.europa.eu/thinktank/en/document.html?reference=EPRS_BRI(2017)599411#:~:text=While%20'capacity%2Dbuilding'%20suggests,change%2C%20borne%20by%20local%20actors.

D4D Meeting – March 23rd		
Topic	**Comments / Feedback**	**Response / Implications**
	established concept that could be included in the framework.	We have integrated the stages/concepts from the Data Value Chain into the analytic framework, in order to enhance it's utility, particularly when applied to D4D projects contemplating Capacity Development as a *Means* or *Process*.
Metrics/Indicators	How can this study be informed/feed into global debates about measurement and indicators? GDB and the SDGs are examples but there are other major Capacity Development indicators that could be incorporated to help to make the work ''mappable''.	Project approval documents, in many cases, do not explicitly speak to/state the indicators or metrics that will be used to evaluate the impact of the project; Explicit consideration of metrics/indicators should be a program design consideration. This has been incorporated into the analytical framework.
Unit of Analysis: Capacity *vs* Capabilities *vs* Competencies	It is often difficult to translate trainings into improved capabilities and institutional capacities. The level of complexity of different Capacity Development initiatives varies significantly across projects. How can the framework be tied to a matrix with a set of related indicators, especially at the outputs and outcomes level, support this?	Identifying the target level for Capacity Development initiatives is an important design consideration. A given project may have more than one target level but it was also noted that it may be important to consider if the individual level is enough for sustainability or should the sector, context or entity also be considered to ensure sustainability. The Unit of Analysis has been incorporated into the analytic framework.
Analytic Framework	The framework is an interesting approach, and it would be interesting to further flesh out outcomes. For example, capabilities is a loaded term with many related concepts.	Incorporating the Data Value Chain into the framework as well as providing a glossary to tease out the various concepts/terms can increase the framework's utility.
Scaling Innovations	Scaling innovations is important for all our work. Thinking about and documenting strategies that work when we design interventions would mean avoiding reinventing the wheel for each project, which would be an important contribution.	The proposed best practices in terms of strategies could be a shared document that is an important input into all project designs. The environmental scans being carried out through parallel initiatives could ultimately contribute to the development of a resource centre to "house"/"direct shared knowledge content.

Table 6: : Feedback & Responses - OD4D Consultation Session - 25-03-2021

OD4D Meeting – March 25th		
Issue	**Comments / Feedback**	**Response / Implications**
Sustainability	Sustainability of Capacity Building is a critical component, but so difficult to achieve. We need to build mechanisms intentionally to ensure the new skills/capacities sustain.	It was suggested that a success factor could be to have institutions embedded/involved in the design of those programs. This is critical to help sustainability. There is also a need to ensure that everyone is on the same page on the issues, updates and innovations so they can feel they are part of the endeavor.
Metrics/Indicators	This must be an important consideration. Every level at activities, outputs, outcomes, should have its own assessment tools to be able to capitalize from learnings.	This was a reoccurring concern throughout the consultations and has been explicitly represented in the framework to ensure that it is a component of project proposals and evaluations.

OD4D Meeting – March 25th		
Scaling Innovations	It would be interesting to better understand the ''who'' of the programs. Who is going to lead the scaling initiatives? Are these efforts most efficient in universities, in public service, etc. The level of specificity at who we target when we design intervention is also important.	In terms of specificity and who should be the target, the unit analysis that has been included as part of the framework forces the researchers to think this through from the outset of the project. One participant pointed out that based on their experience in the government as a public official, a key objective was to ensure Capacity Building of the officials leading the cross-agency work on data and Capacity Building for their own peers. As they felt empowered through this work, they are still involved in its deployment.
Data Value Chain	The Data Value Chain is an interesting concept that could be included in the framework. It helps to identify the ''who/actors'' at different stages in the process.	The stages/concepts from the Data Value Chain have been integrated into the analytic framework, in order to enhance it's association with the existing DVC constructs and its utility.
Demand-side Capacity building	We heard some success factors in past conversations. Identifying patterns would help to lead to success factors (Institutionalisation, sustainability, etc.) The idea of "demand-driven capacity building" is important, as an additional success factor, where Gov agencies, for example, would be genuinely interested to institutionalize such programs across Gov/sectors, etc. So, success is more likely when it is demand driven and not the other way around.	The elements of the Data Value Chain that encourage perception of value; promote data use culture and data-driven behaviour change, all contribute to creating demand. These concepts have been woven into the analytic framework and should be an active part of project design considerations. Project approval documents, in many cases, do not explicitly speak to/state the indicators or metrics that will be used to evaluate the impact of the project; Explicit consideration of metrics/indicators should be a program design consideration. This has been incorporated into the analytical framework.
Outcome Based Design	What is more important than a competency design is an outcome - based design, and then the competencies get developed as you work towards your outcome	This will be a function of whether the project is seeking to develop Capacity as a *Means*, or as the *End* in and of itself. The unit of analysis incorporated into the Framework should require project designers to contemplate this issue.
Sustainability	Sustainability of Capacity Building is a critical component, but so difficult to achieve. There is a need to build mechanisms intentionally to ensure the new skills/capacities sustain.	It was suggested that a success factor could be to have institutions embedded/involved in the design of those programs. This is critical to help sustainability. There is also a need to ensure that everyone is on the same page on the issues, updates and innovations so they can feel they are part of the endeavor.

The discussions and feedback from these internal consultation sessions helped to determine the final elements of the DCD Analytic Framework.