

# Daten für Entwicklung

## Nutzungsperspektiven in der deutschen Entwicklungszusammenarbeit

### Einleitung

„Würde ich als BMZ-Referentin des Querschnittsreferates belastbare Prognosen über Menschenrechtsverletzungen kennen, könnte ich gezielter Vorschläge für deren Schutz im jeweiligen Vorhaben machen.“

„Würde ich als GIZ-Auftragsverantwortlicher der E-Learning-Plattform atingi mehr über die zukünftige Nachfrage auf dem lokalen Arbeitsmarkt wissen, könnte ich bedarfsgerechte Qualifikationen anbieten, um die Jobchancen von Jugendlichen zu erhöhen.“

„Würde ich als KfW-Projektmanager in Mali mehr über den Status eines schwierigen Infrastrukturprojektes in einer aus Sicherheitsgründen unzugänglichen Region erfahren, könnte ich Auszahlungen vorerst zurückhalten, um eine Mittelfehlverwendung zu vermeiden.“

Diese fiktiven Beispiele zeigen welch enormes Potential sich hinter der Nutzung von Daten<sup>1</sup> verbirgt, um unsere alltägliche Arbeit effizienter und wirkungsvoller zu gestalten. Ähnlich wie bei der Erfindung anderer Alltagstechnologien, z.B. der Elektrizität, erfolgt deren breite Nutzung jedoch nicht automatisch, sondern bedarf eines längeren Veränderungsprozesses. So hat das Vorhandensein von Elektrizität nicht unmittelbar zu einer starken Verbreitung und wirtschaftlichen Entwicklung geführt – die geschah erst in einem Prozess über mehrere Jahrzehnte. Auch die systematische Nutzung von Daten in der deutschen EZ steckt aktuell noch in den Kinderschuhen.

Mit diesem Diskussionspapier möchten wir den Prozess anstoßen bzw. beschleunigen. Verlässliche Portfoliodaten sind die wichtigste Grundlage. Aber in der Entwicklungszusammenarbeit stehen immer mehr externe Daten zur Verfügung, seien es Sensordaten, Klimadaten, Satellitendaten, Umfragedaten oder Daten aus digitalen Anwendungen. In einem iterativen nutzerzentrierten Ansatz wollen wir gemeinsam die Chancen einer verantwortungsvollen Datennutzung proaktiv ausschöpfen und die deutsche Entwicklungszusammenarbeit noch stärker auf Wirksamkeit ausrichten. Wir wollen zeigen, dass dies auch ohne die risikobehaftete Nutzung personenbezogener Daten möglich ist und die Vorreiterrolle Europas beim Datenschutz weiter ausbauen.

---

<sup>1</sup> Definition Daten: Daten bezeichnen als Plural von Datum durch Beobachtungen, Messungen oder statistische Erhebungen gewonnene (Zahlen)Werte. Texte, Bilder und Audio/Video-Aufnahmen zählen ebenfalls zu Daten. Aus Daten können, die entsprechenden Kompetenz vorausgesetzt, Informationen gewonnen werden.

## 1. Vision & Ziele

Unsere Vision ist eine datenbasierte Entwicklungszusammenarbeit. Wir wollen Daten systematisch als Grundlage für konkrete Entscheidungsfindungen und eine evidenzbasierte Politikgestaltung nutzen, um den Einsatz unserer begrenzten Ressourcen so effizient und unsere Arbeit so wirksam wie möglich zu gestalten. Damit tragen wir zu einem sicheren und vertrauenswürdigen, aber auch offenen, transparenten und interoperablen Datenökosystem bei, wie es die Bundesregierung im Rahmen ihrer Datenstrategie anstrebt.

Derzeit sammelt das BMZ Portfoliodaten zu EZ-Maßnahmen und greift punktuell auf öffentlich zugängliche Datenbanken von internationalen Organisationen, Forschungseinrichtungen und Think Tanks zu. Die KfW nutzt beispielsweise seit mehreren Jahren Daten für „Remote Management, Monitoring and Verification“ (RMMV). Bei der GIZ entstehen große Datenmengen bei der Nutzung der e-learning-Plattform atingi. Am DEval entstehen Datenkorpora beispielsweise im Rahmen von Text Mining zur Analyse von Dokumenten der Programmplanung, Berichterstattung und Evaluation. Da es bislang jedoch noch keinen systematischen Austausch zur Nutzung externer Daten zwischen BMZ, den Durchführungsorganisationen und den Ressortforschungs-/Evaluierungseinrichtungen gibt, findet eine gemeinsame Analyse von EZ-internen und externen Daten für evidenzbasierte Entscheidungen bislang nur in Einzelfällen statt.

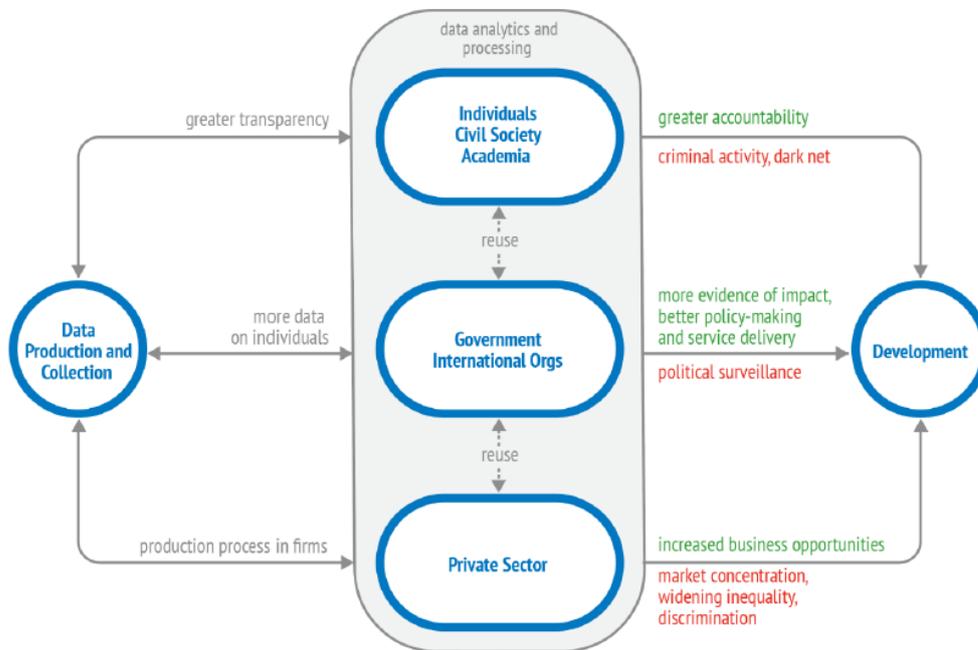
Deshalb streben wir an, die Verfügbarkeit relevanter interner und externer Daten für die deutsche Entwicklungszusammenarbeit zu verbessern. Der Zugang zu geeigneten Datenquellen aus dem BMZ, den Durchführungsorganisationen, den Ressortforschungs-/Evaluierungseinrichtungen, öffentlich zugänglichen Datenbanken oder Datenpartnerschaften<sup>2</sup> soll vereinfacht und ggf. automatisiert werden. Außerdem wollen wir die allgemeine Datenkompetenz der Mitarbeitenden erhöhen und eine Datenkultur in der deutschen Entwicklungszusammenarbeit etablieren.

Wir wollen gemeinsam mit den Durchführungsorganisationen sowie dem DIE und DEval, unsere Kräfte bündeln, wie das der im Entwurf vorliegende World Development Report der Weltbank in nachstehender Grafik zusammenfasst.

---

<sup>2</sup> Datenpartnerschaften bezeichnen Kooperationen zwischen dem BMZ und weiteren Organisationen oder Institutionen, die darauf basieren, dass bestimmte Daten miteinander geteilt werden oder ein gemeinsamer Zugang zu kostenpflichtigen Datenbanken eingerichtet wird. In Abstimmung sind derzeit Datenpartnerschaften mit GSMA und IOM.

Figure O.1. How data can support development: A theory of change



Source: WDR 2021 team.

Note: Positive impacts are shown in green; negative impacts are shown in red.

## 2 Hintergrund: Daten in der EZ

In der Entwicklungszusammenarbeit entstehen immer mehr Daten<sup>3</sup>. Neben „traditionellen“ Datenquellen wie dem Zensus, Haushaltsumfragen oder Verwaltungsdaten können auch aus neuen, „alternativen“ Datenquellen wie beispielsweise Sensor- oder Satellitendaten, Daten aus geografischen Informationssystemen, unstrukturierten Daten (Big Data), wie bspw. der automatisierten Analyse von Dokumenten, sowie Nutzerdaten digitaler Anwendungen eine erhebliche Menge an Informationen generiert werden. Zudem veröffentlichen immer mehr nationale und internationale Institutionen bestimmte Datensätze in Form von Open Data.<sup>4</sup> Daten werden hierdurch zu einer zunehmend wichtigeren Entscheidungsgrundlage für Neuvorhaben

<sup>3</sup> Hier kann zwischen verschiedenen Datenarten unterschieden werden. „Small Data“ bezeichnet die Menge an Daten, die klein genug ist, um mit normalen PC-Anwendungen verarbeitet zu werden (z.B. Portfoliodaten, Governance-Indikatoren, etc.) unterschieden werden. Im Gegensatz dazu beschreibt „Big Data“ besonders große Mengen an Daten, die meist als „digitale Brotkrümel“ anfallen, d.h. es handelt sich um Daten, die durch die Digitalisierung menschlicher (Inter-)Aktionen passiv von Nutzern digitaler Endgeräte/Dienstleitungen „emittiert“ werden (z.B. anonymisierte Nutzerdaten, Sensordaten, Satellitendaten, etc.)

<sup>4</sup> Hierzu zählen u.a. Flucht- und Migrationsdaten (IOM), Klimadaten (DWD), Lebensmittelpreise (FAO), Bevölkerungsdaten (UN), Satellitendaten (NASA/ESA), Governance Indikatoren (Weltbank) sowie die Development Data Library (USAID)

werden, z.B. durch ein besseres Verständnis für Muster wachsender Städte, sich ausbreitender Dürren oder Klimaentwicklungen. Bei laufenden Vorhaben ermöglichen Daten Prognosen bzw. Simulationen und verbessern die Möglichkeiten für Monitoring und Evaluierung.

Die zunehmende „Vermessung der Welt“ in allen Bereichen birgt jedoch auch große Risiken. Die Risiken liegen zum einen im Bereich der Daten selbst, wie zum Beispiel unzureichende Datenqualität, Datenvoreingenommenheit und Herausforderungen im Bereich von Datenschutz und Datensicherheit. Zum anderen bestehen Risiken mit enorm schädlichem Ausmaß durch Cyberbedrohungen und Cyberkriminalität, aber auch in Bezug auf die unlautere Verwendung und Weitergabe der Daten durch die zugriffsberechtigten Personen.

### 3. Rahmenbedingungen

Der Umgang mit Daten in der deutschen Entwicklungszusammenarbeit ist eingebettet in die Datenstrategie der Bundesregierung. Die Durchführungsorganisationen GIZ und KfW entwickeln ebenfalls eigene Datenstrategien. Die in diesem Diskussionspapier vorgeschlagenen Ziele und Maßnahmen sollen die Umsetzung der Datenstrategie der Bundesregierung vorantreiben und einen Handlungsrahmen für die Durchführungsorganisationen schaffen.

#### 3.1 Datenstrategie der Bundesregierung

Die Datenstrategie der Bundesregierung zielt darauf ab, regierungsinterne und externe Daten besser für Entscheidungsprozesse nutzbar zu machen und die Effektivität der öffentlichen Verwaltung durch datengetriebenes Handeln zu erhöhen. Dabei wird der Schaffung von Standards zur Datenqualität und –interoperabilität sowie der Förderung des „open data“ eine besondere Bedeutung beigemessen. Zudem sollen die Ressorts wo erforderlich und sinnvoll Datenzentren einrichten, um Rohdaten zu sammeln und besser zu nutzen. Datenbearbeitungs- und -analysekompetenzen sowie Expertise für die systematische und evidenzbasierte Evaluation politischer Maßnahmen sollen aufgebaut und gebündelt werden. Die vier zentralen Handlungsfelder lauten:

1. Datenbereitstellung verbessern und den Datenzugang sichern: Auf- und Ausbau wettbewerbsfähiger und nachhaltiger Dateninfrastrukturen und eines vertrauenswürdigen Ökosystems aus verschiedenen Akteuren, die auf Grundlage eines gemeinsamen Verständnisses und gemeinschaftlicher Regeln miteinander agieren.
2. Datennutzung steigern: Identifizierung geeigneter Anwendungsfälle und Aufbau eines Datenökosystems zur Nutzung von Innovationspotenzialen.
3. Datenkompetenz erhöhen und Datenkultur etablieren: Konzipierung und Umsetzung geeigneter Fortbildungsmaßnahmen zum Umgang mit Daten und Einstellungskriterien für neuen Mitarbeiterinnen und Mitarbeiter
4. Den Staat zum Vorreiter machen: Der Staat muss bestehende rechtliche Rahmenvorgaben effektiv durchsetzen sowie durch gute Rahmenbedingungen und geeignete Sicherheitsarchitekturen Vertrauen schaffen. Gleichzeitig muss der Staat auch seine eigenen Daten in geeigneter Form als open data zur Verfügung stellen.

Um diese Handlungsfelder für die Entwicklungspolitik umzusetzen sollen die Ressorts u.a. folgende Maßnahmen implementieren:

- Einrichtung interner Datenlabore: Alle Bundesministerien und/oder ihre nachgeordneten Behörden sollen in eigener Verantwortung interne Datenlabore etablieren. Das BMZ ist hier als Vorreiterbehörde genannt.
- Forschungsdatenzentren<sup>5</sup>: Die Bundesministerien werden selbst oder in nachgelagerten Einrichtungen Forschungsdatenzentren errichten bzw. aufbauen, sofern diese fachlich sinnvoll erscheinen. Diese sind Ansprechpartner für die Nutzung von (nicht eingestuft) Daten. Das BMZ als Vorreiterbehörde soll sich hier bei der Setzung von Standards einbringen.
- Open-Data-Koordinatoren: In allen Behörden der Bundesverwaltung sollen Koordinatoren und Koordinatorinnen eingeführt werden, die als Ansprechpartner für alle Belange rund um das Bereitstellen von Daten fungieren können.
- Chief Data Scientist: Es wird angestrebt, dass alle Bundesministerien ein/e Chief Data Scientist oder vergleichbare Rolle einführen. Der/die Chief Data Scientist würde mit einem Kernteam von Daten-Analystinnen und Analysten zusammenarbeiten.

### 3.2 Datenstrategien der Durchführungsorganisationen

Die GIZ arbeitet seit August 2020 an einer Datenstrategie. Der Fokus liegt auf der Zusammenführung interner Projekt- und Portfoliodaten und Weiterentwicklung der Datenkompetenz von Mitarbeitenden zur Pflege und Analyse dieser Daten. Im Frühjahr 2021 soll ein erster Entwurf der GIZ-Datenstrategie vorliegen, der bis Ende 2021 finalisiert wird.

Die KfW entwickelt seit Mitte 2020 eine Digitalisierungsstrategie, welche eine Datenstrategie als Kapitel beinhaltet. Ziel ist es, datenbasiertes Arbeiten auf allen Ebenen in der KfW zu implementieren. Um dieses Ziel zu verwirklichen, strebt die KfW die Nutzung organisationsinterner und -externer Daten sowie von open data zur Projektvorbereitung, -steuerung, -monitoring und -evaluierung an.

### 3.3 Datenstrategien von Ressortforschungs-/Evaluierungseinrichtungen

Am DEval wird momentan eine Projektleitung für das Management von Evaluierungs-/Forschungsdaten besetzt. Aufgabe dieser Stelle wird u.a. die Erarbeitung einer Datenstrategie sein. Darüber hinaus befasst sich das DEval in seinen Evaluierungen neben qualitativen Daten und quantitativen Umfragedaten u.a. mit Geodaten, Text Mining/Machine Learning sowie Social-Media-Daten.

## 4. Prinzipien für die Sammlung, Analyse und Nutzung von Daten in der deutschen EZ

Bei der Entwicklung der Prinzipien orientieren wir uns an den Principles for Digital Development (insb. Design with the User, Use Open Data, Address privacy and security, Be collaborative) sowie den bereits vorliegenden Datenstrategien der Vereinten Nationen, der EU und der Bundesregierung. Gemeinsame Prinzipien schaffen Rechts- und Verfahrenssicherheit und ermöglichen Synergieeffekte.

---

<sup>5</sup> S. z.B.: Rat für Sozial- und Wirtschaftsdaten (RatSWD) bzw. KonsortSWD (<https://www.konsortswd.de/>)

## 4.1 Grundprinzipien

### 1. Gewinnbringend („Design with the User“):

Nur wenn die bereitgestellten Daten einen Mehrwert an Informationen für eine relevante Problemstellung mitbringen, werden diese auch genutzt. Eine Anhäufung nicht genutzter Daten ist kontraproduktiv. Um letzteres zu vermeiden und Daten bedarfsgerecht zur Verfügung zu stellen arbeiten wir eng mit den Endnutzern zusammen.

### 2. Open by Default („Use Open Data“).

Daten, die seitens des BMZ und der Durchführungsorganisationen bei der Erfüllung ihrer öffentlich-rechtlichen Aufgaben erhoben werden, sollen standardmäßig der Öffentlichkeit bereitgestellt werden. Dies umfasst u.a. Daten, die bei der Vorbereitung, Umsetzung, und Evaluierung von EZ-Maßnahmen anfallen. Damit schafft die EZ ein öffentliches Gut und befolgt die Vorgaben von EU und Bundesregierung („public money, public data“). Außerdem kann somit eine Duplizierung des Aufwands zwischen den Akteuren der deutschen EZ vermieden werden. Ausnahmen gelten bspw. in Bezug auf Datenschutz, Betriebs- und Geschäftsgeheimnisse und Schutz des geistigen Eigentums und Entscheidungsbildungsprozesses. Dieser Abwägungsprozess muss auf ausformulierten Normen bzw. Prinzipien beruhen und standardisiert angewendet werden und darf nicht subjektiven Beurteilungen überlassen werden.

### 3. Aufstellung eines Risikomanagements („Address privacy and security“).

Um Daten vor den in Kapitel 1 genannten potenziellen Risiken wie Cyberkriminalität oder Datenmissbrauch zu schützen, muss eine Strategie zum präventiven und reaktiven Risikomanagement aufgestellt werden. Komponenten sollten u.a. sein:

- DSGVO-konformer Schutz der Daten;
- Erhöhung der Datenkompetenz der Zugriffsberechtigten;
- Geeignete Infrastruktur zum Schutz vor Cyberkriminalität;
- Planung einer uniformen Datenarchitektur zum Schutz vor verminderter Datenqualität.

### 4. DSGVO-Konformität

Alle Daten müssen nach EU-Verordnung DSGVO-konform gesammelt und genutzt werden. Personenbezogene Daten oder organisationsinterne Daten (z.B. mitarbeiterbezogene Daten sowie Daten, die ausschließlich interne Prozesse betreffen) dürfen nicht gemeinsam genutzt oder veröffentlicht werden.

## 4.2 Anforderungen an die Datenbereitstellung

- Die Bereitstellung der Daten muss, ebenso wie die Datensammlung, bestimmten Prinzipien unterliegen. Die Basis für diese Prinzipien bilden die anerkannten „FAIR“-Prinzipien<sup>6</sup> für Datenmanagement: **F**indable, **A**ccessible, **I**nteroperable und **R**eusable. Findable (Auffindbarkeit): Die gesuchten Daten sollten sowohl für Menschen als auch für Computer leicht zu finden sein.

---

<sup>6</sup> S. z.B.: [https://www.konsortswd.de/wp-content/uploads/RatSWD\\_WP\\_274.pdf](https://www.konsortswd.de/wp-content/uploads/RatSWD_WP_274.pdf)

- **Accessible (Zugänglichkeit):** Daten und Metadaten sollten langfristigarchiviert und unter definierten Voraussetzungen – d.h. unter Beachtung des Datenschutzes, der Vertraulichkeit und ethischer Überlegungen – so verfügbar gemacht werden, dass sie für zugangsberechtigte Nutzende leicht und jederzeit abgerufen und heruntergeladen oder lokal genutzt werden können. Die Nutzerfreundlichkeit steht dabei im Vordergrund.
- **Interoperable (Interoperabilität):** Die Daten sollten so vorliegen, dass sie ausgetauscht, interpretiert und mit anderen Datensätzen kombiniert werden können. Die Metadaten für organisationsübergreifende und offene Daten müssen dafür einheitlich sein.
- **Reusable (Wiederverwendbarkeit):** Das Ziel ist es, die Wiederverwendung von Daten zu optimieren. Um dies zu erreichen, sollten Metadaten und Daten gut beschrieben sein, damit sie in verschiedenen Umgebungen repliziert und/oder kombiniert werden können.

Ergänzend sind folgende Prinzipien wichtig:

- **Klassifizierung:** Systematische Erfassung und Klassifizierung der Datenquellen hinsichtlich der Anwendungsfelder (Vorbereitung, Steuerung, Monitoring und Evaluierung von Projekten / Vorhaben) und Sektoren.
- **Auswertungstools:** Systematische Erfassung und Bewertung unterschiedlicher Tools zur datenquellenunabhängigen Datenauswertung und –darstellung

### 4.3 Anforderungen an die Datenqualität

Die Qualität der bereitgestellten Daten ist von zentraler Bedeutung und muss stets so hoch wie möglich gehalten werden. Nur wenn diese gegeben ist, sind die daraus gezogenen Informationen bzw. Rückschlüsse belastbar. Um die Qualität herzustellen und zu wahren, müssen klar definierte Verantwortlichkeiten festgelegt werden und die Daten sollten folgende Dimensionen erfüllen;



- **Konsistent:** Die Information, abgelegt an Ort A, muss übereinstimmen mit den relevanten Daten, die an Ort B abgelegt sind.
- **Einheitlich:** Die Daten sollten fortlaufend auf dieselbe Art und Weise, d.h. in den korrekten gemeinsam vereinbarten Regeln folgenden Formaten abgespeichert und abgebildet werden.
- **Objektiv:** Daten müssen objektiv, das heißt streng sachlich und wertfrei sein.
- **Aktuell:** Die Daten sollten stets aktualisiert werden und auch entsprechend gekennzeichnet werden.

- **Komplett:** Die Daten sollten weitestgehend die komplette Lage im jeweiligen Bereich darstellen.
- **Präzise:** Die Informationen sollten die Realität so gut wie möglich widerspiegeln.

#### 4.4 Synergien zwischen BMZ und Durchführungsorganisationen nutzen („Be collaborative“)

Damit die gemeinsame Aufbereitung und Nutzung von Daten durch BMZ und Durchführungsorganisationen erfolgreich wird, ist eine gute Abstimmung zwischen den Akteuren wichtig:

- **Datenformate, Schnittstellen und Metadaten:** Um die Nutzbarkeit und Interoperabilität der Daten sicherzustellen ist ein essenzielles Ziel die Datenformate sowie Schnittstellen und Metadaten zu vereinheitlichen/standardisieren sowie eine API festzulegen. Ein solches System umfasst auch Überlegungen zum Datenzugang gestaffelt nach Zugangsrechten und Datenklassifizierung.
- **Analysetools:** Auf Basis der Vereinheitlichung der Datenformate, Schnittstellen und Metadaten sollen Empfehlungen für priorisierte IKT-Bausteine für Analysetools erstellt werden. Bisher kommt z.B. im BMZ IBM Cognos zum Einsatz, bei der GIZ Microsoft PowerBI, bei DEval Stata, R und SPSS sowie Python-basierte Machine-Learning/Text-Mining-Systeme. Das erschwert den Austausch und die Zusammenarbeit an gemeinsamen Analyseprojekten.
- **Fortbildungen:** Fortbildungen sollten organisationsübergreifend gemeinsam und verpflichtend sein, um Datenkompetenz nachhaltig aufzubauen
- **Einstellungspolitik:** Die Durchführungsorganisationen und Ressortforschungs-/Evaluierungseinrichtungen sollten genauso wie das BMZ nachhaltig und langfristig qualifiziertes Personal aufbauen (z.B. durch die Einstellung von Data Scientists, o.a. geeigneten Personen zur Auswertung / Analyse bzgl. spezifischer übergeordneter Fragestellungen im bi- und multilateralen Bereich des BMZ)

## 5. Mögliche strategische Weichenstellungen in der laufenden Legislaturperiode

### 5.1 Stärkung der Datenkompetenz

Die Datenkompetenz in den Partnerländern und in Deutschland sollte als wichtiges strategisches Ziel weiter ausgebaut werden. Die aktuelle Studie von Qlik und Accenture „The Human Impact of Data Literacy“ zeigt, dass immer noch eine große Lücke herrscht zwischen dem, was mit datengestützten Entscheidungen erreicht werden könnte und dem, wie häufig und wie fachgerecht Mitarbeitende mit Daten umgehen (können). Data literacy, also die Fähigkeit, Daten zu lesen, zu bearbeiten und zu analysieren ist nicht ausreichend vorhanden. So bewertet ein Drittel aller Befragten den Umgang mit Daten am Arbeitsplatz als überfordernd. Das führt bei einem Drittel der Befragten sogar dazu, dass sie das Erledigen solcher Aufgaben komplett vermeiden. Diese Studie macht deutlich, dass neben der technischen Umsetzung der Datensammlung und –verarbeitung, insbesondere die Ausbildung der Datenkompetenz des Personals von elementarer Wichtigkeit für ein datengetriebenes Arbeiten ist.

Zur Erweiterung der Datenkompetenz würden sich beispielsweise bedarfsgerechte, an das jeweilige Datenkompetenzniveau angepasste Fortbildungen für alle Mitarbeitenden des BMZ,

der Durchführungsorganisationen und der Ressortforschungs-/Evaluierungseinrichtungen anbieten. Darüber hinaus ist die gezielte Personalrekrutierung von Personen, die sowohl fachliche als auch Datenkenntnisse aufweisen, ein wichtiger Weg zur Steigerung der Datenkompetenz der Mitarbeiterschaft insgesamt. Datenkompetenz sollte entsprechend in angemessener Weise als Auswahlkriterium berücksichtigt werden.

## 5.2 Datenkultur

Für die erfolgreiche Umsetzung sollte eine gemeinsame Datenkultur im BMZ für alle Mitarbeitenden in allen Stufen der Hierarchie geschaffen werden. Die Erhöhung der Datenkompetenz aller Mitarbeiter\*innen bildet hierfür die Grundlage. Weitere Elemente der Datenkultur sind:

- Schutz persönlicher Daten: Einbettung von "privacy by design" in alle Datenpraktiken. Einholung einer informierten Zustimmung und Gewährleistung der Anonymisierung der Daten vor der Veröffentlichung.
- Einhaltung der höchsten ethischen Standards: Verankerung der Datenpraxis in der UN-Charta und im internationalen Menschenrechtsrahmen. Sicherstellung, dass Datenprozesse und Partnerschaften dem öffentlichen Wohl dienen.
- Verantwortlichkeit der Datenverwaltung: Effektive Datenverwaltung und -steuerung praktizieren, um eine hohe Datenqualität, Sicherheit und Verantwortlichkeit in Übereinstimmung mit den einschlägigen institutionellen Richtlinien und Vorschriften zu gewährleisten.
- Die Grenzen von Daten erweitern: Erforschung neuer Praktiken und innovativer Technologien zur Verbesserung der Datenverfügbarkeit und zur Erweiterung der Erfassung unterrepräsentierter Bevölkerungsgruppen in Partnerländern durch Datenerhebung und -disaggregation.
- Sensibilisierung für die Grenzen des datenbasierten Arbeitens: Bewertung von Lücken, Risiken und Verzerrungen bei der Verwendung von Daten<sup>7</sup>. Blinde Flecken und mögliche negative Auswirkungen der Datennutzung aktiv hinterfragen.

## 5.3 Einrichtung einer Community of Practice for Data Driven Development Cooperation

Im Rahmen einer Community of Practice (CoP) sollte pilothaft die gemeinsame Nutzung externer Daten für Entwicklung erprobt werden. Basierend auf einem gemeinsamen Grundverständnis über Prinzipien der Datennutzung entsprechend der Rollen der beteiligten Organisationen könnte die CoP folgende, aufeinander aufbauende Aktivitäten durchführen:

1. Identifizierung mehrerer Anwendungsfälle und Dokumentation von entwicklungspolitisch relevanten Problemstellungen, die durch datenbasiertes Arbeiten adressiert werden könnten
2. Erörterung von datenbasierten Indikatoren in Wirkungsmatrizen
3. Vorschläge für Design Thinking Sprints im BMZ digilab
4. Ggf. Aufbau einer Testplattform für relevante gegenseitig nutzbare Daten mit APIs für niedrigschwelligen Zugang zu benötigten externen Daten

---

<sup>7</sup> S. z.B.: <https://www.unwomen.org/-/media/headquarters/attachments/sections/library/publications/2018/gender-equality-and-big-data-en.pdf?la=en&vs=3955>

5. Vereinheitlichung/Standardisierung von Datenformaten, Schnittstellen und Metadaten sowie der Festlegung einer API. Überlegungen zum Datenzugang, Zugriffsrechten und Datenklassifizierung.
  6. Entwicklung eines Risikomanagementsystem
  7. Austausch über die Nutzung von KI/Machine Learning für big data
  8. Austausch mit Organisationen in unseren Partnerländern
  9. Austausch mit Forschungsdatenzentren anderer Ressorts/Bereiche sowie entsprechenden Dachorganisationen (z.B. KonsortSWD) zur Identifikation guter Praktiken des Datenmanagements und der Datenbereitstellung
  10. Untersuchung der Möglichkeiten der Integration von Portfolio- und Projektdaten (z.B. in Maßnahmen erhobene Daten zum Zwecke des Monitoring, Evaluierung, Forschung) in das Datenmanagement.
- 
-

## 6. Anhang

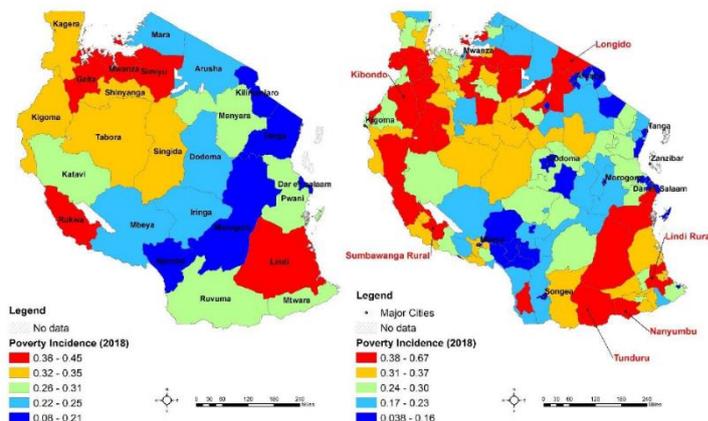
### 6.1 Anwendungsbeispiele für die Kombination verschiedener Datenquellen

#### Armutsinzidenz Tansania:

Der Fall Tansania liefert eine gute Illustration für die Vorteile von einer gemeinsamen Analyse verschiedener Datensätze. Mit den Daten einer Haushaltsbefragung konnte die Regierung das Ausmaß der Armut in nur 20 Bezirken des tansanischen Festlandes erfassen. Durch die Kombination der Haushaltsumfragedaten mit Satellitenbilddaten wurde es möglich, die Armut für jeden der 169 Bezirke Tansanias zu schätzen. Durch die Kombination der beiden Datenquellen konnte die Auflösung des Armutsbildes um das Achtfache erhöht werden, ohne dass die Genauigkeit darunter litt. Dieses Beispiel zeigt, dass es einen hohen Mehrwert bietet, wenn verschiedene Datensätze gemeinsam analysiert werden. Damit die Datensätze optimal und effizient analysiert werden können, ist ein gemeinsames Portal, auf dem sie vorliegen, von großem Vorteil. So wird sichergestellt, dass die Datensätze automatisch aktualisiert sind und so stets in aktueller Form in einem Format vorliegen, dass eine gemeinsame Analyse möglich ist.

Map O.2. Combining satellite imagery with household survey data increases the resolution of the poverty map of Tanzania

a. Poverty map using the Household Budget Survey (20 districts)      b. Poverty map combining the data in panel a with satellite imagery (169 districts)



#### Covid-19 Pandemie:

Die Covid-19 Pandemie hat die EZ in vielen Partnerländern verändert und neue Maßnahmenplanung notwendig gemacht. Die bisherigen Covid-19-Erfahrungen haben gezeigt, wie öffentliche Statistiken, die aus Daten des privaten Sektors erstellt wurden - Kreditkartenausgaben, Beschäftigung und Unternehmenseinnahmen - ein neues Werkzeug für die politische Analyse darstellen. Die über Mobiltelefone gesammelten Daten, wie z. B. Call Detail Records (CDR) und GPS-Standortdaten, ermöglichten die Messung der Bevölkerungsdichte, der Reisemuster und der Durchmischung der Bevölkerung in Echtzeit und mit hoher Auflösung. Das verbesserte die Ausrichtung der politischen Maßnahmen und die Erstellung von epidemiologischen Modellen. Wenn relevante Daten aus dem öffentlichen und privaten Sektor zusammen mit Portfolio-daten der EZ-Organisationen automatisch in einem Portal aggregiert werden, können sie für genau solche politischen Maßnahmen in Partnerländern zusammen analysiert und somit effizient genutzt werden.

#### Estland X-Road:

In Estland hat die Regierung das nationale Datensystem X-Road eingerichtet, um die persönlichen Daten der Bürger für die Nutzung durch Behörden und teilnehmende Unternehmen sicher zu verwalten. X-Road ist eine Open-Source-Lösung für den Datenaustausch auf verschiedenen Ebenen, die es verknüpften öffentlichen und privaten Datenbanken ermöglicht, automatisch Informationen auszutauschen und dabei Vertraulichkeit, Integrität und Interoperabilität zu gewährleisten. Bei diesem System müssen die Daten nur einmal an Behörden und teilnehmende Unternehmen übermittelt werden. Sie werden dann automatisch an andere teilnehmende Stellen weitergeleitet. Die bisherige Auswertung von X-Road zeigt, dass durch die geschaffene Transparenz Vertrauen aller Akteure zueinander erhöht wird. Durch die nationale Reichweite, der einfachen Bedienung und der Vollständigkeit der Daten, wird Gleichberechtigung zudem aller Akteure gefördert. Die Kultur des Datasharings auf einem gemeinsamen Tool wurde hier mit großem Erfolg und Mehrwert für alle Beteiligten umgesetzt.

## 6.2 Beispiele für Datenquellen bei anderen Organisationen

- IOM: Flucht- und Migrationsdaten
- DWD: Klimadaten
- FAO: Lebensmittelpreise
- GSMA: Mobilfunkdaten, Mobile Payment Daten
- UN: Bevölkerungszahlen
- NASA, ESA: Satellitendaten
- Weltbank: Governance Indikatoren
- USAID: Development Data Library (DLL); Alle quantitativen Daten, die von USAID oder von einer von USAID unter Vertrag genommenen oder geförderten Organisation erhoben werden, müssen auf der DDL hochgeladen und gespeichert werden.

## 6.3 BMZ-Projekte als mögliche Datenquellen

- atingi: Die digitale Lernplattform des BMZ (atingi) ermöglicht Menschen im globalen Süden jederzeit und überall kostenlosen Zugriff auf hochwertige digitale Lerninhalte. Die Plattform generiert DSGVO-konforme Nutzerdaten, die wertvolle Einblicke in verschiedene Lebensbereiche darstellen können. So kann es beispielsweise für die Planung eines Bildungsprojekts sehr sinnvoll sein, anonymisiert den Bildungsstand von Nutzergruppen und das Interesse an bestimmten beruflichen Ausbildungen in einer Region oder einem Land zusammen mit anderen relevanten Daten von bisherigen Projekten oder open-source Daten zum Thema Armut zu analysieren.
- DataCipation: Einrichtung einer interaktiven Informationsplattform zur Verbesserung des Dialogs zwischen Bürgern und Bürgerinnen und politischen Vertretern und Vertreterinnen der AU und ihrer Mitgliedsstaaten
- Earth Observation for Sustainable Development Programm: Die kostenfreien Satellitenbilder und Dienste der ESA besser in Nutzung bringen
- digilab: Deutsche Entwicklungszusammenarbeit u.a. in der Datennutzung digitaler und effizienter gestalten. Ziel ist es, diese mit agilen Methoden wie Design Thinking, effektiver effizienter und moderner zu gestalten.
- Fair Forward: Globale Maßnahme der Bundesregierung (BMZ) und verfolgt u.a. das Ziel lokale KI-Kompetenzen in Afrika und Asien auszubauen.

- IKT Bausteine: Plattformen für modernes E-Government. Entwicklung einer internationalen Plattform mit standardisierten und spezifizierten IKT Bausteinen (u.a. Messaging, E-Payment, Registration, Authentication, Data Exchange, Data Validation) als globales öffentliches Gut. Darüber hinaus wird eine Arbeitsgruppe eingerichtet, die sich mit Kernstandards befasst, die für alle Komponenten gelten. Diese Standards werden sich mit Daten- und Interoperabilitätsanforderungen befassen, die für alle Bausteine benötigt werden.