

Why Data Matters

By Heather Leson, Data Literacy Lead



International Federation of Red Cross and Red Crescent Societies
Fédération internationale des Sociétés de la Croix-Rouge et du Croissant-Rouge
Federación Internacional de Sociedades de la Cruz Roja y de la Media Luna Roja
الاتحاد الدولي لجمعيات الصليب الأحمر والهلال الأحمر

DATA PLAYBOOK: SLIDEDECK 4



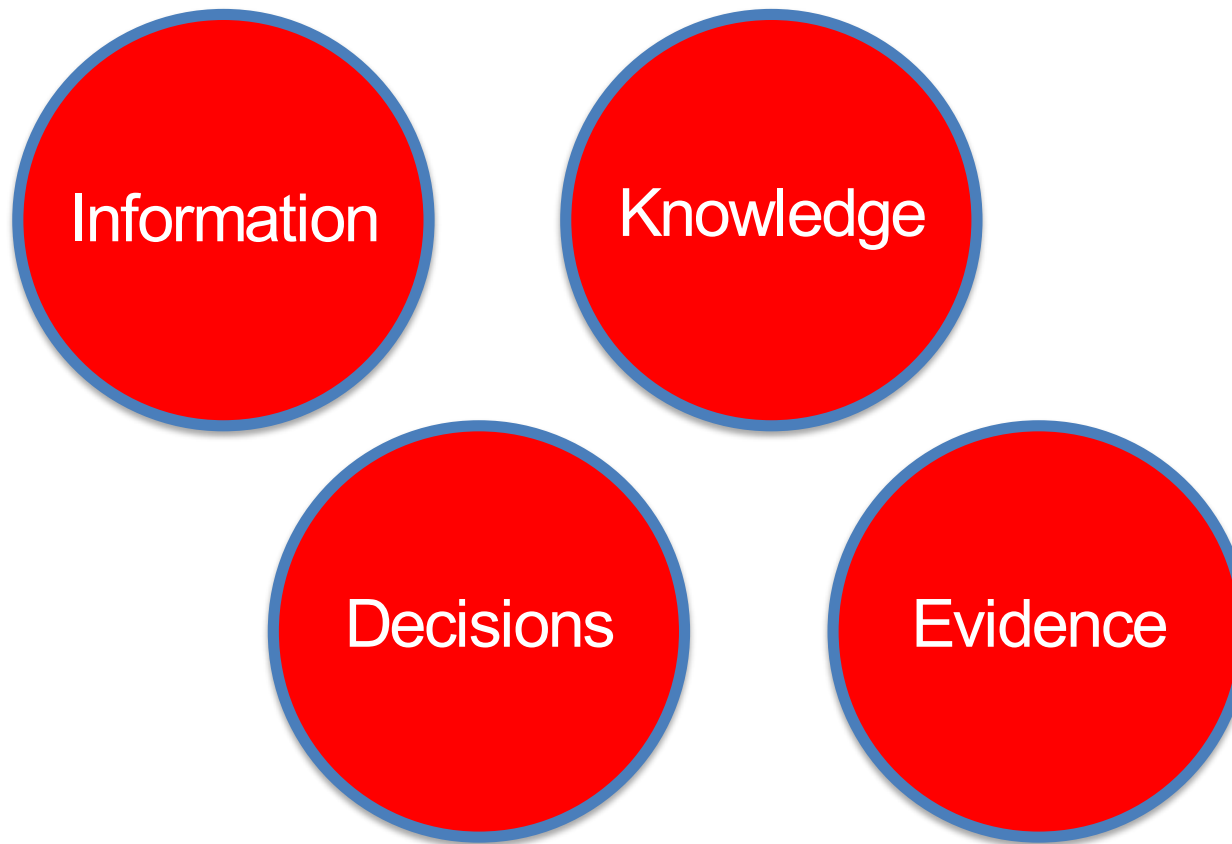
The Data Revolution is here. Are we Data Ready?



International Federation of Red Cross and Red Crescent Societies
Fédération internationale des Sociétés de la Croix-Rouge et du Croissant-Rouge
Federación Internacional de Sociedades de la Cruz Roja y de la Media Luna Roja
الاتحاد الدولي لجمعيات الصليب الأحمر والهلال الأحمر

DATA PLAYBOOK: SLIDEDECK 4

Data can lead to:



IFRC Global Operations (GO) snapshot (June 2018)



International Federation of Red Cross and Red Crescent Societies
Fédération internationale des Sociétés de la Croix-Rouge et du Croissant-Rouge
Federación Internacional de Sociedades de la Cruz Roja y de la Media Luna Roja
الاتحاد الدولي لجمعيات الصليب الأحمر والهلال الأحمر

Data is part of our Leadership



IFRC is the Secretariat, National Societies, and volunteers.

We aim to be a data-driven organization making evidence-based decisions. It is cited in our 2020 strategy.

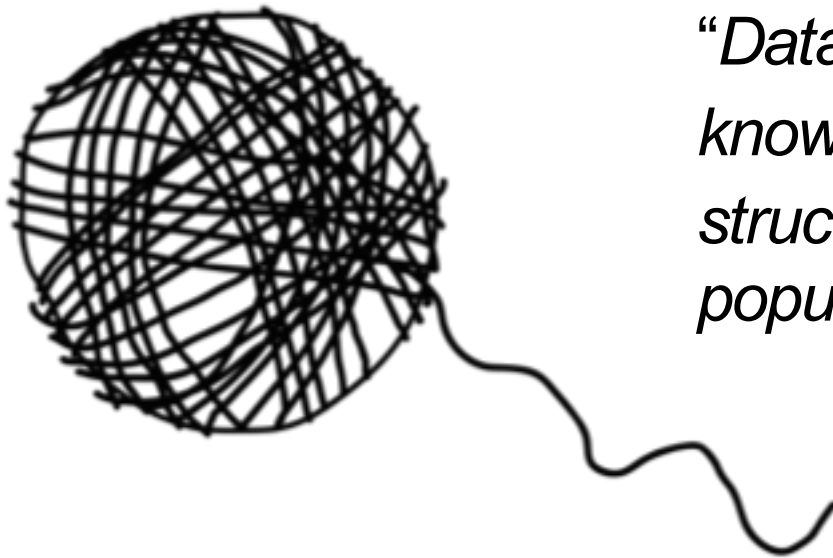
Data-literate is not the same as data-skilled



*“A data-literate organisation is one that shares a culture of data and a strong vision of the future. Most people invested in this vision will have no analytic interaction with data and may never need to.”**

*Source: Open Data Institute

What is Data Literacy?



*“Data Literacy includes the skills, knowledge, attitudes, and social structures required for different populations to use data.”**

How can we build an ecosystem of data ready colleagues?

*Source: School of Data

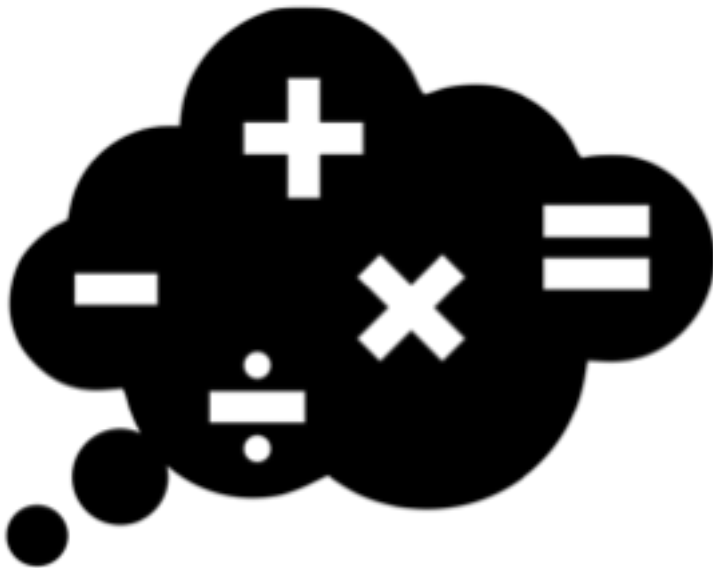


International Federation of Red Cross and Red Crescent Societies
Fédération internationale des Sociétés de la Croix-Rouge et du Croissant-Rouge
Federación Internacional de Sociedades de la Cruz Roja y de la Media Luna Roja
الاتحاد الدولي لجمعيات الصليب الأحمر والهلال الأحمر

What does data literacy mean for me?

Role	Task
IM/Operations/PMER/Health	Deliver projects with information products/Assess project and programme delivery
Marketing Communications	Excellent data/analysis, narrative for storytelling, Brand and fundraising
IT	Assess and support data products/tools, provide infrastructure
Training	Provide e-learning, workshops and technical training
Manager	Strategic planning, staff development, organization development
Community served	Provide data, obtain help/services, get feedback

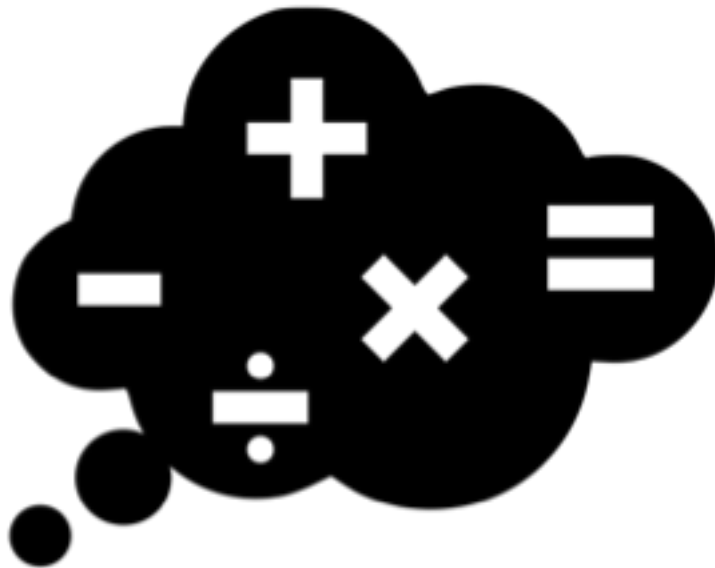
Potential benefits of focusing on Data literacy



- Teamwork / Collaboration
- Increased Accountability/Transparency
- Organizational Effectiveness (reuse, decrease of duplication)
- Financial improvements
- Competencies / Skills



How can we prove “Data Readiness?”



We measure many things at IFRC.
How Might Data Readiness
measurements be incorporated into
existing frameworks:

- PMER/MEAL
- Surge/IM
- ICT Health Check/Digital Divide
- OCAC/BOCA
- Program Planning
- Competencies



Data Pipeline

When we talk about “data”, people often focus on the **skills**, **tools** and the **process** steps for delivery of data products like a “dataset.”

The ‘Data Pipeline’* is an example of data ready skills. We all have varying levels of know-how.

*Source: School of Data



International Federation of Red Cross and Red Crescent Societies
Fédération internationale des Sociétés de la Croix-Rouge et du Croissant-Rouge
Federación Internacional de Sociedades de la Cruz Roja y de la Media Luna Roja
الاتحاد الدولي لجمعيات الصليب الأحمر والهلال الأحمر

Graphic Source: School of Data

DATA PLAYBOOK: SLIDEDECK 4

Humanitarian Data Teams: Supporting Skills

Humanitarian Business

- Cluster coordination
- Assessments
- Operational planning
- Logistics/Roster Management
- Disaster Risk Reduction
- Response preparedness
- Disaster relief/Recovery
- Thematic Areas of Focus
- Health, Gender and Social Inclusion



Network

- Clients
- Humanitarian agencies
- Development agencies
- Access to skilled people, information managers, database managers, data analysts
- Businesses
- Investors, sponsors and donors

Business Skills

- Leadership
- Strategic business planning
- Marketing & Sales
- Customer relations
- People management & HR
- Administration
- Public speaking
- Problem resolution
- Finance and accounting skills
- Delegating tasks
- Motivating team

Soft Skills

- Strategic, proactive, creative, innovative and collaborative
- Curious about data
- Influence without authority
- Problem solver
- Hacker /Maker mindset

Humanitarian Data Teams: Technical Skills

Math and Statistics

- Machine learning
- Statistical modeling
- Supervised learning & Unsupervised learning
- Statistical computing (e.g. R)
- Relational algebra



Programming

- Computer science fundamentals
- Scripting language (i.e. Python, javascript)
- Filtering scripts (i.e. D3.js)
- Web development
- Experience with xaaS like AWS

Data Management

- Data modelling
- Data collection
- Data refinement and cleaning
- Database, SQL and NOSQL
- Parallel databases and parallel processing
- Open Data standards
- API's
- Hadoop and Hive/Pig

Information Management

- GIS & Mapping
- Survey methodology
- Data analysis
- Finding & using datasets

Communications and Visualization

- Story telling skills
- Translate data-driven insights into decisions and actions
- Interactive dashboards
- Infographics
- Visual art design
- Knowledge of visualisation tools like Tableau, Adobe toolkit



International Federation of Red Cross and Red Crescent Societies
Fédération internationale des Sociétés de la Croix-Rouge et du Croissant-Rouge
Federación Internacional de Sociedades de la Cruz Roja y de la Media Luna Roja
الاتحاد الدولي لجمعيات الصليب الأحمر والهلال الأحمر

Emergency types by region

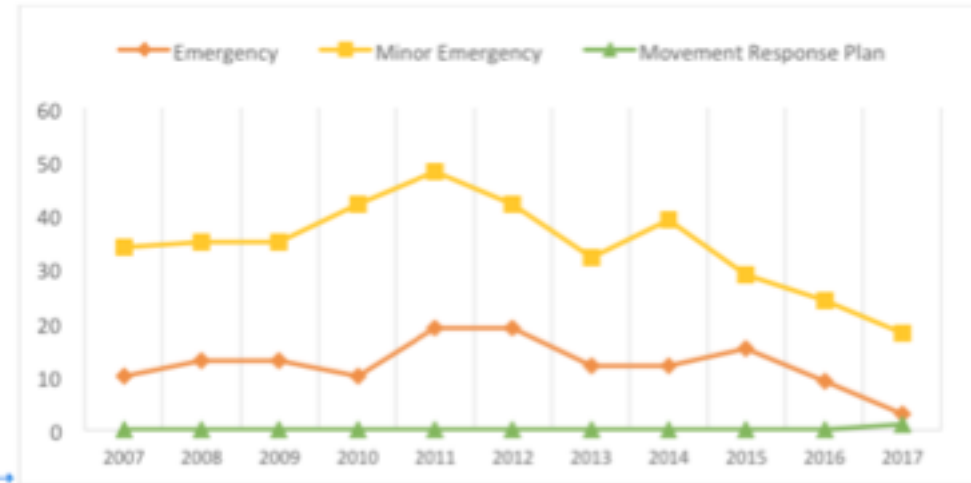
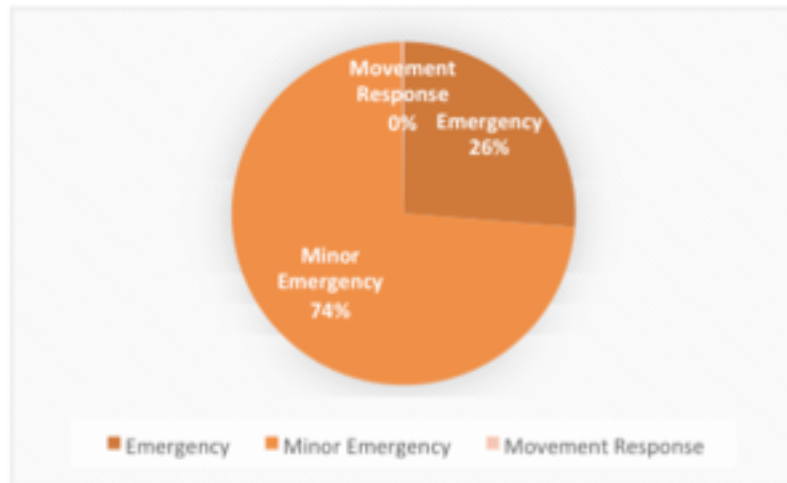
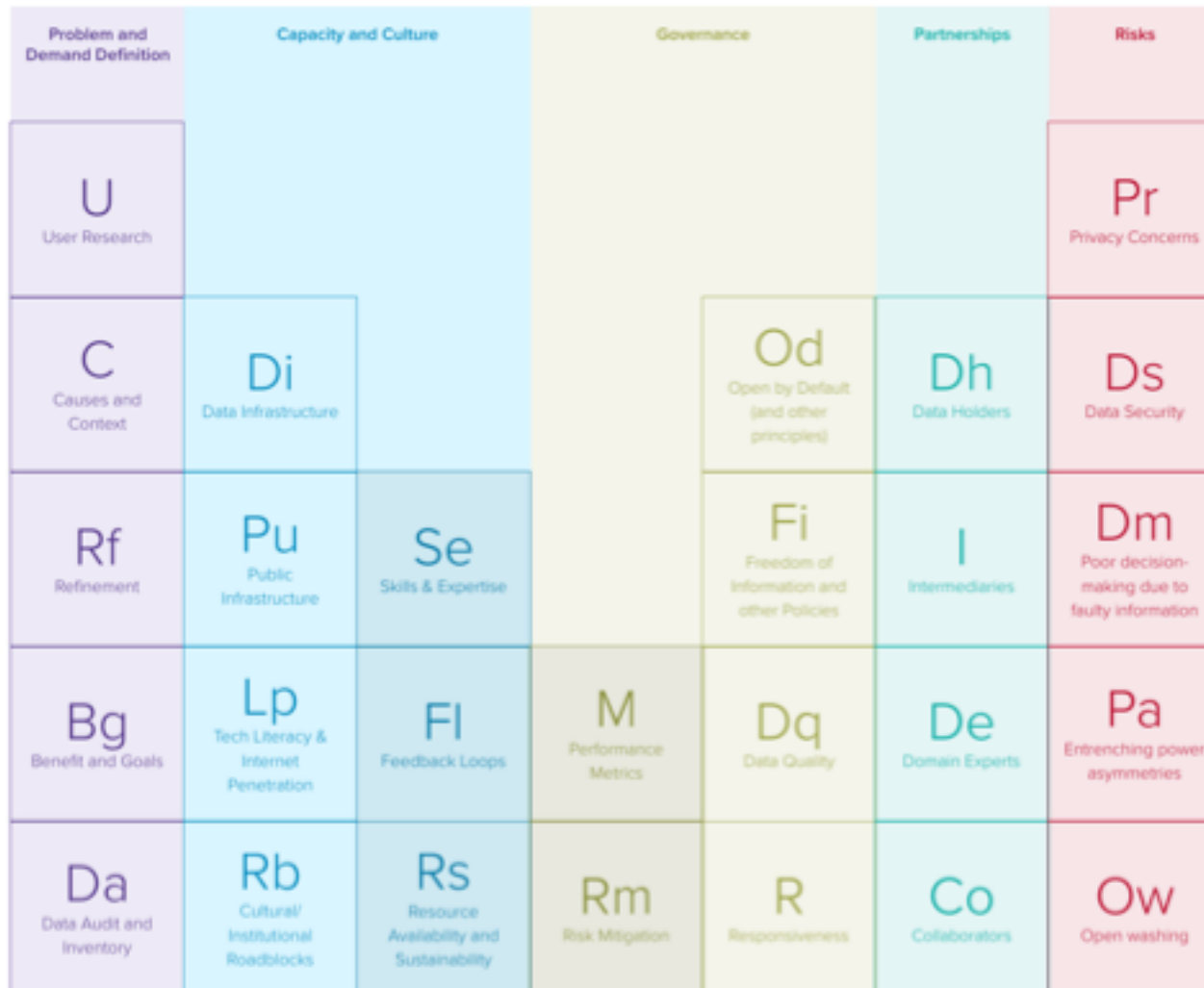


Fig1: Types of Emergencies in Africa by severity between 2007-2017 → **Fig 2: Trends in emergencies in Africa between 2007-2017**

Considering Data Workflows



Data Maturity Framework

Data and Tech Readiness Scorecard

Category	Area	Lagging	Basic	Advanced	Leading
How is Data Stored	Accessibility	Only accessible within the application where it is collected	Can be accessible outside the application but proprietary format, requiring specialized analysis software	All machine readable in standard open format (CSV, JSON, XML, database)	All machine readable in standard open format and available through an API
	Storage	Paper	PDFs or Images	Text Files	Databases
	Integration	Data sits in the source systems	Data is exported occasionally and integrated in ad hoc manner	Central data warehouse - realtime aggregation and linking (Automatic)	External data also integrated
What is Collected?	Relevance and Sufficiency	The data you are collecting on subjects of interest is irrelevant to the problem you want to solve; ie you want to do predict which students need extra support to graduate on-time but don't have data on graduation outcomes	Some of the data you have is relevant, but it is insufficient because key fields are missing, ie no data on academic behavior or attendance history, etc.	You have data that is helpful and relevant for solving the problem but not sufficient to solve it well, ie you have yearly academic and demographic information but are missing extra-curricular activities, or interventions they were targeted with	You have all the relevant data about all the entities being analyzed and it's sufficient to solve the problem you are tackling
	Quality	Missing rows (people/address level entities missing in the data)	Missing columns (variables missing)	No missing data but errors in data collection such as typos	No missing data and no errors in data collection
	Collection Frequency	Once and never again	yearly	frequently	realtime
	Granularity	City level aggregates	Zipcode/Block level aggregates	Individual level (person or address) level data	Incident/Event level data
	History	No History Kept - old data is deleted	Historical data is stored but updates overwrite existing data	Historical data is stored and new data gets appended with timestamp, preserving old values	All history is kept and new data schema gets mapped to old schema so older data can be used
Other	Privacy	No privacy policy in place	no PII can be used for anything	ad-hoc approval process in place that allows selected PII data to be used for selected/approved projects	Software defined/controlled privacy protection that allows analytics to be done while preserving privacy based on predefined policies
	Documentation	no digital documentation or metadata: data exists but field descriptions or coded variables are not documented	data dictionary exists (variables and categories defined)	data dictionary plus full metadata available (including conditions under which the data were captured)	data dictionary plus full metadata available including collection assumptions, what's not collected, and potential biases



Data Literacy Menu

1. Connect

- Informal Data Working Group
- Data Stories
- Ecosystem Map
- Data Simulations

2. Learn

- Build on existing curriculum
- Connect with other Data Literacy Organizations
- Sessions
- Excel around the world

3. Create

- IFRC Data Playbook:
- Templates, checklists, best practices, scenarios and recipes.

4. Measure & Impact

- Responsible Data Policy (in draft)
- Revise IT Policies
- Data Readiness Metrics/KPIs/Competencies



THANK YOU

Heather Leson

heather.leson@ifrc.org

@heatherleson

skype: heatherleson



International Federation of Red Cross and Red Crescent Societies
Fédération internationale des Sociétés de la Croix-Rouge et du Croissant-Rouge
Federación Internacional de Sociedades de la Cruz Roja y de la Media Luna Roja
الاتحاد الدولي لجمعيات الصليب الأحمر والهلال الأحمر

DATA PLAYBOOK: SLIDEDECK 4