Why Data Matters

By Heather Leson, Data Literacy Lead
The Data Revolution is here.
Are we Data Ready?
Data can lead to:

- Information
- Knowledge
- Decisions
- Evidence
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
## What does data literacy mean for me?

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

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

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

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

**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

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

**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

DATA PLAYBOOK: SLIDEDECK 4
Emergency types by region

**Fig 1:** Types of Emergencies in Africa by severity between 2007-2017

**Fig 2:** Trends in emergencies in Africa between 2007-2017
Considering Data Workflows
<table>
<thead>
<tr>
<th>Category</th>
<th>Area</th>
<th>Lagging</th>
<th>Basic</th>
<th>Advanced</th>
<th>Leading</th>
</tr>
</thead>
<tbody>
<tr>
<td>Accessibility</td>
<td>Only accessible within the application where it is collected</td>
<td>Can be accessible outside the application but proprietary format, requiring specialized analysis software</td>
<td>All machine readable in standard open format (CSV, JSON, XML, database)</td>
<td>All machine readable in standard open format and available through an API</td>
<td></td>
</tr>
<tr>
<td>Storage</td>
<td>Paper</td>
<td>PDFs or Images</td>
<td>Text Files</td>
<td>Databases</td>
<td></td>
</tr>
<tr>
<td>Integration</td>
<td>Data sits in the source systems</td>
<td>Data is exported occasionally and integrated in ad hoc manner</td>
<td>Central data warehouse - real-time aggregation and linking (Automatic)</td>
<td>External data also integrated</td>
<td></td>
</tr>
<tr>
<td>Relevance and Sufficiency</td>
<td>The data you are collecting on subjects of interest is irrelevant to the problem you want to solve: if you want to do predict which students need extra support to graduate on-time but don’t have data on graduation outcomes</td>
<td>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.</td>
<td>You have data that is helpful and relevant for solving the problem but not sufficient to solve it well. If you have yearly academic and demographic information but are missing extra-curricular activities, or interventions they were targeted with</td>
<td>You have all the relevant data about all the entities being analyzed and it’s sufficient to solve the problem you are tackling</td>
<td></td>
</tr>
<tr>
<td>Quality</td>
<td>Missing rows (people/address level entities missing in the data)</td>
<td>Missing columns (variables missing)</td>
<td>No missing data but errors in data collection such as types</td>
<td>No missing data and no errors in data collection</td>
<td></td>
</tr>
<tr>
<td>Collection Frequency</td>
<td>Once and never again</td>
<td>yearly</td>
<td>frequently</td>
<td>real-time</td>
<td></td>
</tr>
<tr>
<td>Granularity</td>
<td>City level aggregates</td>
<td>Zipcode/Block level aggregates</td>
<td>Individual level (person or address) level data</td>
<td>Incident/Event level data</td>
<td></td>
</tr>
<tr>
<td>History</td>
<td>No History Kept - old data is deleted</td>
<td>Historical data is stored but updates overwrite existing data</td>
<td>Historical data is stored and new data gets appended with timestamp, preserving old values</td>
<td>All history is kept and new data schema gets mapped to old schema so older data can be used</td>
<td></td>
</tr>
</tbody>
</table>

**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