Fit-for-purpose geospatial data for crises: integrating EO and AI

Dr. Nathaniel K. Newlands
(Summerland, British Columbia, Canada)

.

Contributors:
Christy Caudill (Carleton University), Cheila Cullen (City University of New York), Markus Enenkel (World Bank), Rui Kotani (GEO Secretariat)

(Pre-recorded virtual presentation):
2nd Expert WG on Policy and Legal Frameworks for Geospatial Information Management, United Nations Statistics Division (UNSD)

About me…

Sustainable Ecosystems (Food, Water, Energy)

Predictive Analytics, Data Science

- Research Scientist
  Summerland (British Columbia
  Agriculture and Agri-Food Canada)
- Adjunct Assoc. Professor 2014-
  (Geography, University of Victoria/UVic)
- Adjunct Professor 2008-2016
  (Statistics, University of British Columbia/UBC)

International Relations in Statistics,
American Statistical Association (ASA)

President-Elect, The International Environmetrics
Society (TIES)

Deputy-Chair, UN-FAO Global Expert Working Group on
Ecosystem/Nature-based Solutions

Co-chair of Group on Earth Observation (GEO)'s
Disaster Risk Reduction (DDR) Working Group

Podcast: AI and agriculture:
https://trustmakers.ca/ownthescience/2021/11/05/
ais-take-up-farming/

Web:
https://profils-profiles.science.gc.ca/en/profile/dr-nathaniel-k-newlands
https://loop.frontiersin.org/people/125828/overview

Social media:
Twitter: @nnewlands
LinkedIn: nathaniel-newlands-b1105230
Climate change AI: https://community.climatechange.ai/home

Email: nathaniel.newlands@agr.gc.ca

Harnessing the predictive power of information

"A transition from linear to exponential growth of human knowledge has taken place"

YOU ARE HERE — 6th wave

5th wave

4th wave

3rd wave

2nd wave

1st wave
- Iron
- Water power
- Mechanisation
- Textiles
- Commerce

Steam power, Railroad, Steel, Cotton

Electricity, Chemicals, Internal combustion engine

Petrochemicals, Electronics, Aviation, Space

Digital Networks, Biotechnology, Software, Information technology

1785 — 1845 — 1900 — 1950 — 1990 — 2020

Innovation
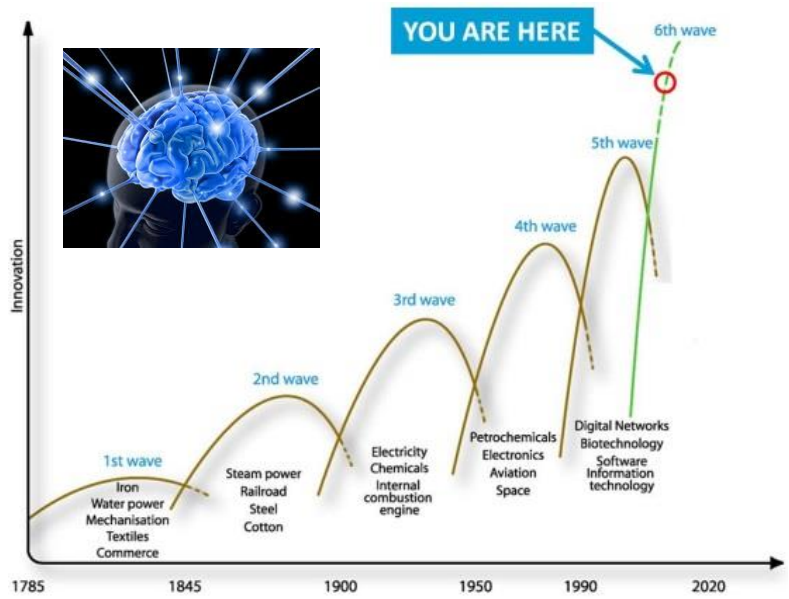
1940s, knowledge doubled every century
1990s, doubled every 25 yrs
Currently doubling every 13 months
Soon, every 12 hrs?

90% of the world's data today was generated during the past 2 years, with 2.5 quintillion bytes of data (10^{18}) added each day…

Google DeepMind

Policy network — $p_{\sigma/\rho}(a|s)$

Value network — $v_\theta(s')$

s — s'
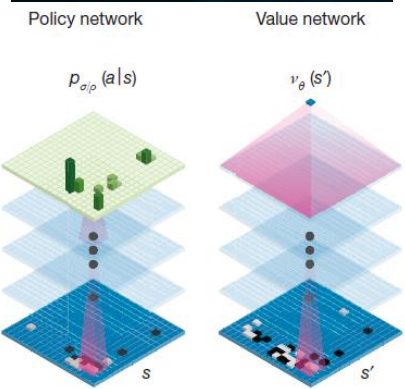
AlphaGo: value/policy networks

IBM

Watson: Deep-learning cloud-based/cognitive apps/APIs

Open, geospatial data, infrastructure

Precision agriculture (automated data collection, analytics)

Enhanced decision making and foresight?

- Complex changes, difficult to anticipate, understand, or plan ahead (needs differ by immediacy and safety risk)
- Cascading, systemic risks: climate change, extreme weather events, disease and pest outbreaks, infrastructure, human health
- Early Warning for All (EW4All) of World Meteorological Organization (WMO); GEO is an implementing partner

Broad use applications, stakeholders
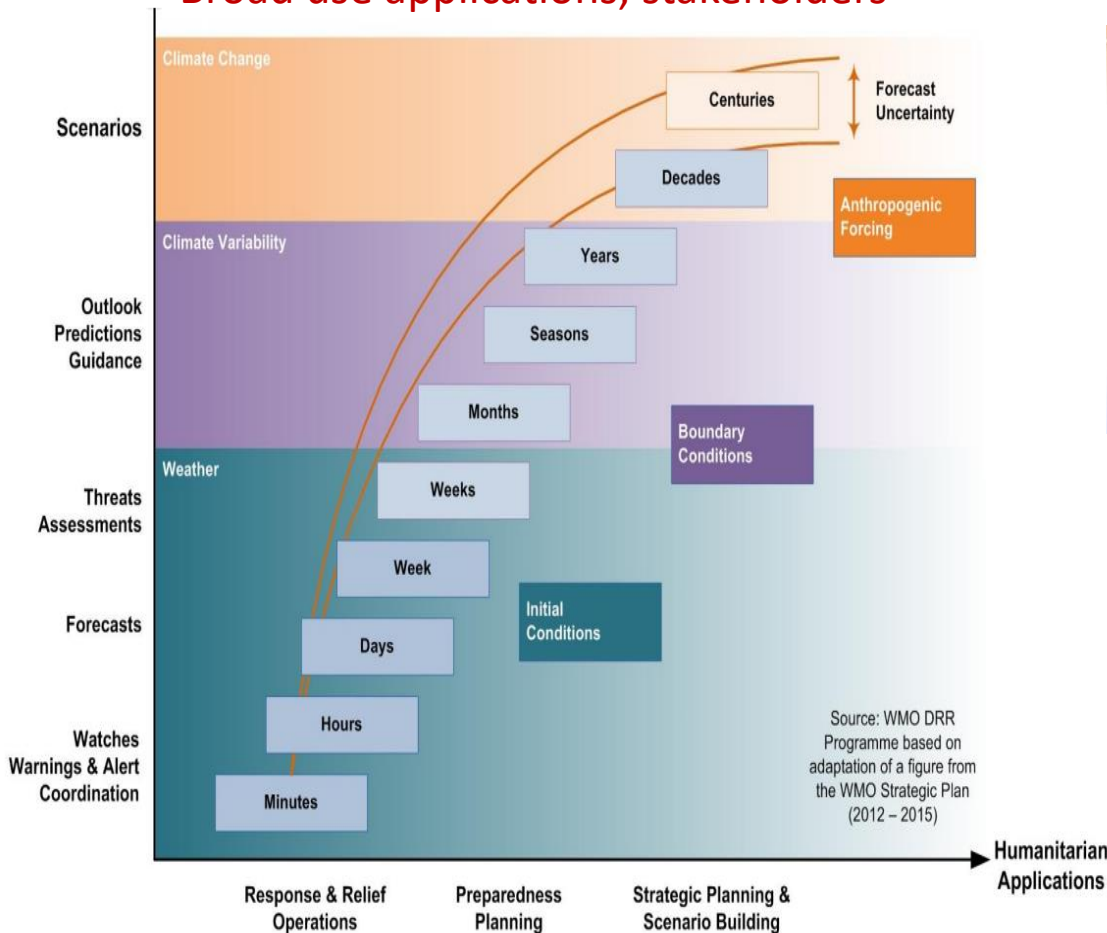
Private-public risk sharing
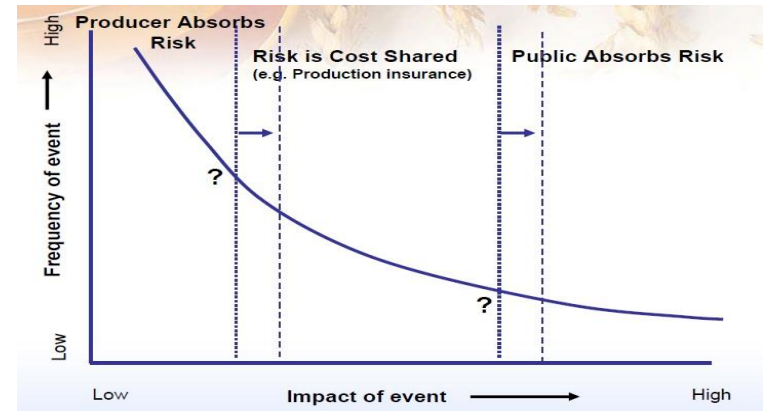
Multi-hazard, vulnerability, impact data

Fit-for-purpose

- designating data as authoritative (trusted, official) is strongly dependent on the intended use of the data and ; multiple purpose(s) may differ between users, with new use-cases evolving over time with technological, AI/ML innovations (training/validation), big data (machine-readable)

- difficult to identify EO authoritative data among the proliferation of data, and/or best integration for research vs. operational use
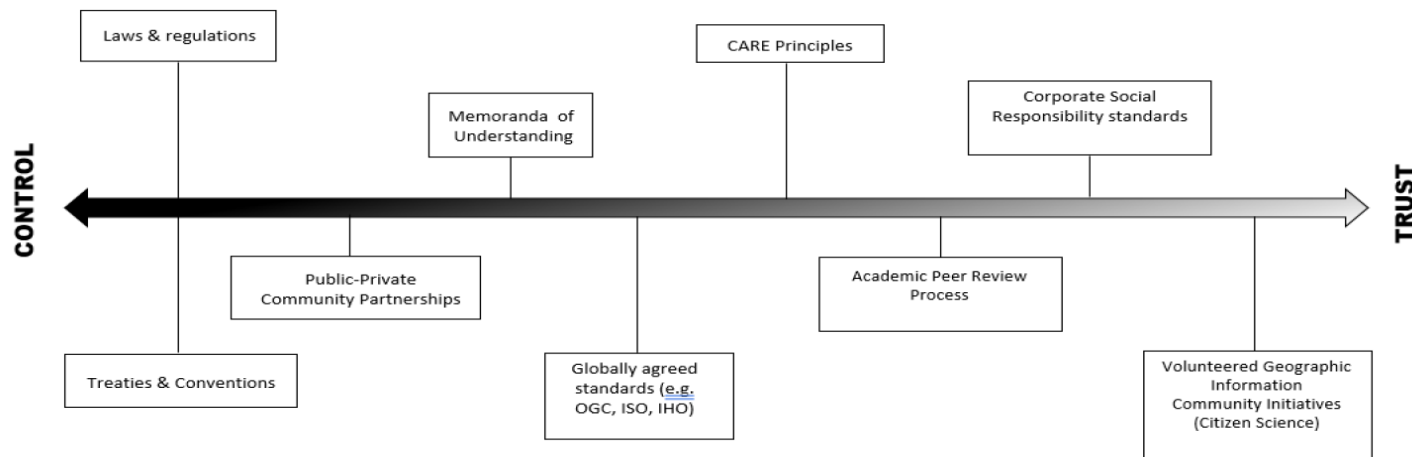
Figure 1. Continuum of control versus trust seen in authoritative data governance.

Key operational rules
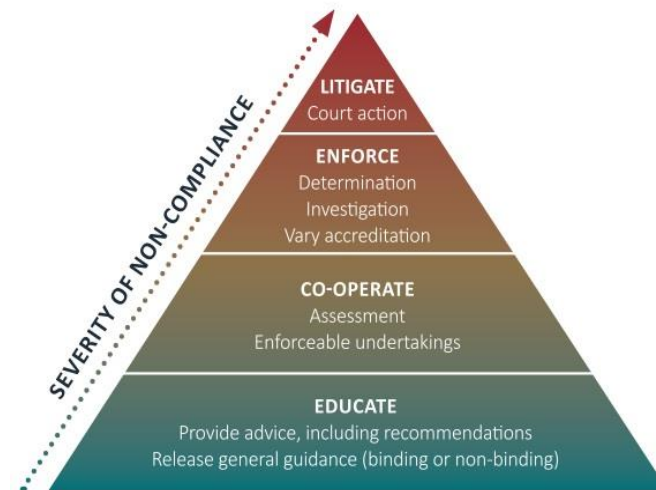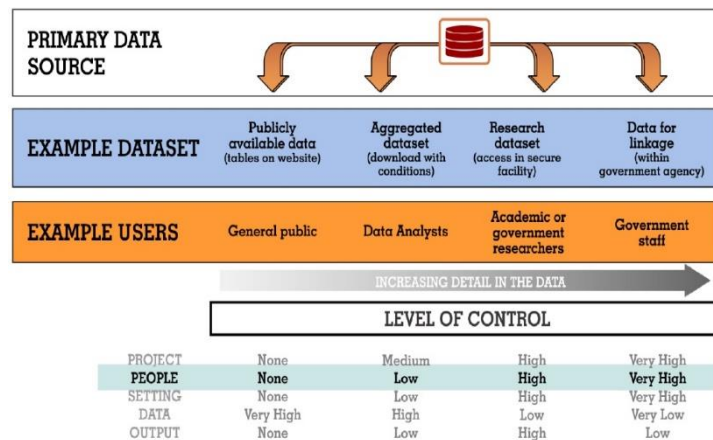
Safeguards to data access?
Data custodians: maintain responsibility for the data they collect
Accredited Data Authorities: technical expertise which could assist data custodians
Trusted users: accredited after demonstrating safeguards for data handling

Streamlined (mandatory) data sharing and release agreements:
(Projects, People, Settings, Data, Output)

PRIMARY DATA SOURCE				
EXAMPLE DATASET	Publicly available data (tables on website)	Aggregated dataset (download with conditions)	Research dataset (access in secure facility)	Data for linkage (within government agency)
EXAMPLE USERS	General public	Data Analysts	Academic or government researchers	Government staff
	INCREASING DETAIL IN THE DATA			
LEVEL OF CONTROL				
PROJECT	None	Medium	High	Very High
PEOPLE	None	Low	High	Very High
SETTING	None	Low	High	Very High
DATA	Very High	High	Low	Very Low
OUTPUT	None	Low	High	Low

SEVERITY OF NON-COMPLIANCE

LITIGATE
Court action

ENFORCE
Determination
Investigation
Vary accreditation

CO-OPERATE
Assessment
Enforceable undertakings

EDUCATE
Provide advice, including recommendations
Release general guidance (binding or non-binding)

Directive on Automated Decision-Making (April 2019): to ensure that Automated Decision Systems are deployed in a manner that reduces risks to Canadians and federal institutions, and leads to more efficient, accurate, consistent, and interpretable decisions made pursuant to Canadian law
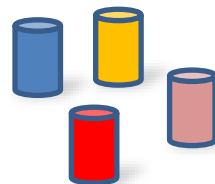
Conceptual 'big data' framework

Strategic:
Identifying long-term or overall aims and means of achieving them

Framework:
Structure, linkage, integration of business processes

Legacy distributed data assets

- Digital/non-digital types
- High fragmentation
- Infrastructure failing
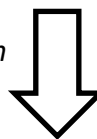- Software risks
- Single points of failure

Public Dissemination, Stakeholder, User Feedback

Public catalogues

- Autonomous data catalogues
- National and Global digital ecosystems

Transition   Transform

Data capture

- Multiple streams
- Heterogeneous
- Structured/Unstructured
- Fit-for-purpose
- Sourcing
- Coherency
- Consistency
- Quality

Ingest

Verification

Digital ecosystem (management, brokering)

- Infrastructure
- Governance (principles, clarity, consistency)
- Technical standards
- Flexibility: Interchangeable data managers
- Improved collaboration using data
- Maintenance, updating, integration, sharing
- Cleansing, preparation
- Points of truth

Prioritization Contingencies Re-evaluation
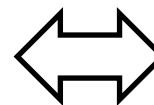
Automation Validation

Business Intelligence (BI) reporting, marketing

- Whole sector challenges
- Value proposition
- Short vs. long-term
- Boosted productivity performance
- Faster adoption, actions,
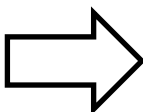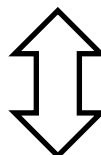- Agile actions, responses, Commercialization

Changing data sourcing
New observing system data
Data reshaping
Testing new algorithms
Piloting/co-learning

Reuse   Remix

Referencing, Archiving

Analytics

- Machine learning, deep learning
- Feature detection, explanation
- Benchmarking, comparability/compatibility
- Algorithms
- Automation
- Validation

Government-wide Digital library

- Assigning identifiers
- Archiving

Group on Earth Observation (GEO)

- **Intergovernmental partnership** (hosted by WMO): 114 governments, 140 inter-governmental, international and regional organizations, 19 private sector and civil society organizations and thousands of world-leading scientists

- promotes **open, coordinated, sustained data sharing and infrastructure** for better research, policy making, decisions, actions

- **GEO Disaster Risk Reduction WG** created to develop and implement a coherent and cross-cutting approach to advance use of EO data in support of national DRR and resilience efforts

- **engaging and collaborating** with all stakeholders, processes, identifying resources to support EO DRR initiatives

- **GEO post-2025 Strategy (Earth Intelligence) – fit-for-purpose operating model:**

  global partnership where data providers and users from all communities work together, leading to better coordination, greater inclusion, reduced duplication, and faster action

Addressing EO Data licensing …

- **Data licensing** has emerged as a factor that is holding back data sharing - complex End User License Agreements, include restrictions on use and require close legal review

- **2013 G8 Open Data Charter (ODC)** (170 governments, organizations) - policies and practices that enable governments and central statistical offices (CSO)s to collect, share, and use well-governed data, to respond effectively, accountably

  Open by Default;
  Timely and Comprehensive;
  Accessible and Useable;
  Comparable and Interoperable;
  For Improved Governance and Citizen Engagement
  For Inclusive Development and Innovation

- **EO data licensing guidance** (GEO Data WG, 2023) promoting the use of standard open data licenses to tackle legal barriers and uncertainty currently inhibiting open data usage (based on FAIR Principles, Open Science initiatives, and the GEO Statement on Open Knowledge).

- Three types of standard open data license (if no restrictions exist, CARE Indigenous Data Governance) should be one of:
  Creative Commons Zero 1.0 Universal Public Domain Dedication (CC0)
  Open Data Commons Public Domain Dedication and License (PDDL) v1.0
  Creative Commons Attribution 4.0 International (CC BY 4.0)

Policy and legal barriers, enablers

- Data quality management – governance strategies, operational systems (long vs. short-term)

- Limited access to free/affordable, historical / near-real time (NRT) "high-resolution" data

- Use-case uncertainty and data limitations (e.g. spatial resolution, dense vegetation, coastal areas, topography)

- Innovation/IT infrastructure (automation vs. humans-in-the-loop; AI and Quantum-ready)

- Lack of personnel, technical capacity to use and process EO data (reliability, privacy)

- Users expect access to a ready-to-use product, transformed from the raw EO data into usable information (inclusive/equitable, participatory co-design for verification)

- Tools and infrastructure needed to process and handle such data can be beyond the capacity of some government entities

- Political and economic context, bureaucracy

Cerbaro M, Morse S, Murphy R, Lynch J, Griffiths G. Challenges in Using Earth Observation (EO) Data to Support Environmental Management in Brazil. *Sustainability*. 2020; 12(24):10411. https://doi.org/10.3390/su122410411

INTERNATIONAL CHARTER SPACE & MAJOR DISASTERS

A worldwide collaboration through which satellite data are made available for the benefit of disaster management

- *initiated by ESA and Centre National d'Etudes Spatiales (CNES)*
- *operational since November 2000*
- *17 members include Canadian Space Agency (CSA)*

Mechanism to make critical space assets available to communities affected by disasters, with EO data free of charge enabling coordination of relief efforts

- Provides critical assistance and intervention to national emergency government agencies by external parties

- Sovereignty, autonomy – users initiate requests, not data providers (automated AI algorithms?)

- Partners, associated bodies, experts could serve as a decision-making body where operations are severely impacted (will and consent)

- Activation is limited to the urgent disaster situation (< 10 d)
- Response phase only (not mitigation, preparedness, recovery)

- Storms, floods, landslides, wildfires and more…
- Policy-linked intervention: systemic vs. sector-specific impact assessment (context relevant) (e.g., agriculture) (privacy safeguards, verification for law enforcement)
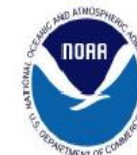
From Big Data to Better Decisions:
DRR and Climate Resilience for Small Island Developing States (SIDS)

Earth Observations for Improving Disaster Risk - Jamaica

- Global Partnerships
- Scientific Assessment
- Funding
- Digital Development
- Education and Training
- Infrastructure
- Checks and Balances
- Capacity Building
- National Disaster Management
- Response Groups
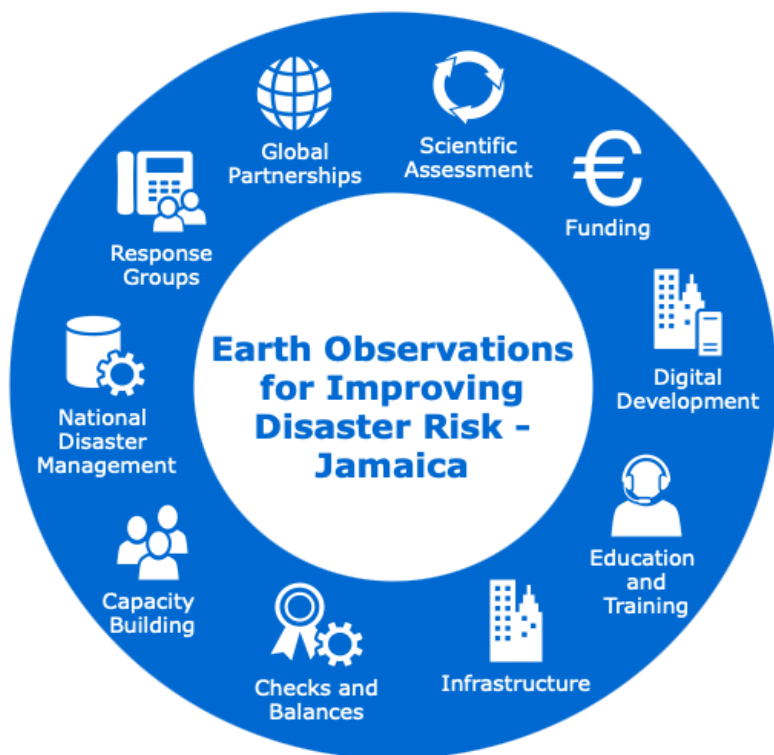
Virtually no contribution to climate change, but highly vulnerable (65 million people, < 1% GHG's)

challenges with regards to equity and inclusion of vulnerable communities

need for bottom-up community engagement, addressing knowledge gaps, data needs empowering youth, capacity-building, co-learning

UN Office for Disaster Risk Reductions (UNDRR) has been calling for integrated approach to address increasingly **systemic nature of disaster risk where events overlap and interplay with multiple risk drivers**

GEO DDR WG is teaming up with stakeholders in Jamaica / Barbados / Caribbean to pilot DRR EO initiatives

fit-for-purpose needs, address multi-sectoral systemic/cascading risks, future scale-up

Thank you!

For more information, please contact:
nathaniel.newlands@agr.gc.ca