An interoperability strategy for SEEA data & models + Next steps for ARIES for SEEA

Ferdinando Villa, Ken Bagstad, Stefano Balbi, Alessio Bulckaen & the ARIES Team

UNCEEA – June 2021

Interoperability: A challenge to the SEEA community

"Without interoperability, data are FAR from FAIR"

– François Soulard

Integration & human nature

“Integration frameworks are like toothbrushes: everyone wants one, but nobody wants to use someone else’s”

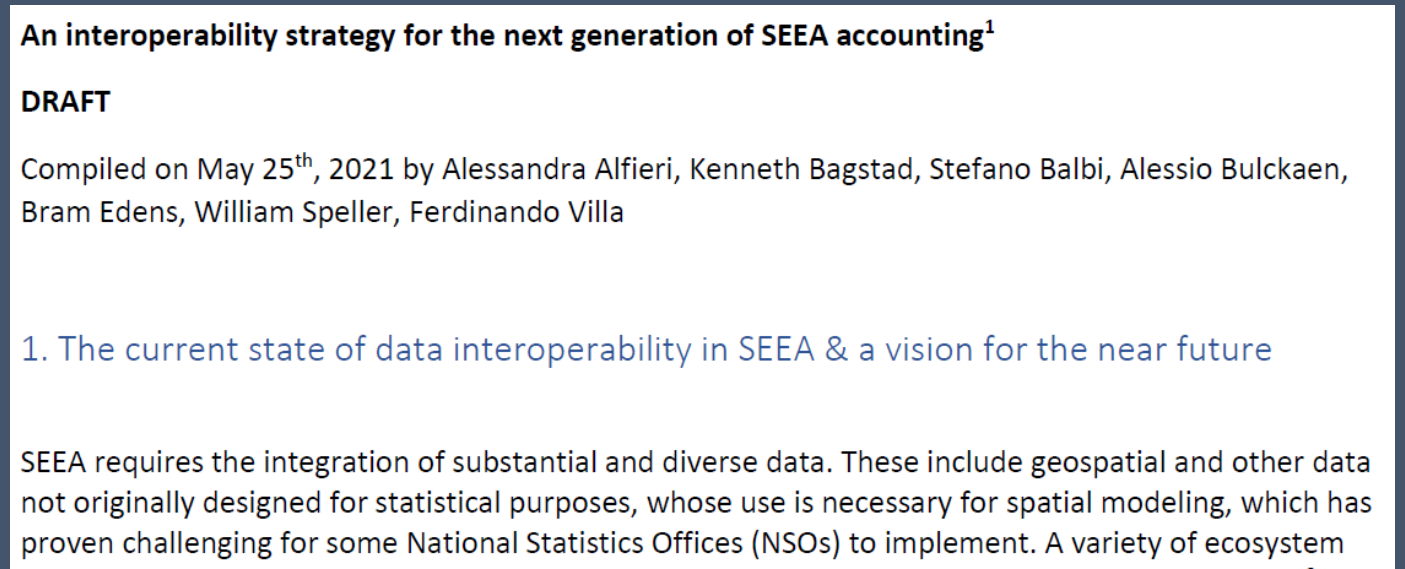
– Jim Boyd, Resources for the Future

A SEEA interoperability & reuse crisis?

- Scientists from the Global North can do great (& painstaking) one-off studies
- We don't know another way - this is how we were trained in grad school;
- Capacity development in the Global South remains very time consuming;
- A fundamental equity issue: those of us on this call (Global North experts) are at a huge advantage. Put yourselves in the place of a junior NSO employee tasked with developing their nation's first SEEA EA accounts. How do we maximize their opportunity for success?
- Also a significant problem for repeated production of accounts everywhere

SEEA interoperability strategy

1. Current state of interoperability & vision for the future
2. Roles & responsibilities (data providers, modelers, institutions incl. NSOs)
3. Implementing the strategy (pilot testing, engaging key stakeholders, governance, training/capacity building)
4. Conclusions



An interoperability strategy for the next generation of SEEA accounting[1]

DRAFT

Compiled on May 25th, 2021 by Alessandra Alfieri, Kenneth Bagstad, Stefano Balbi, Alessio Bulckaen, Bram Edens, William Speller, Ferdinando Villa

1. The current state of data interoperability in SEEA & a vision for the near future

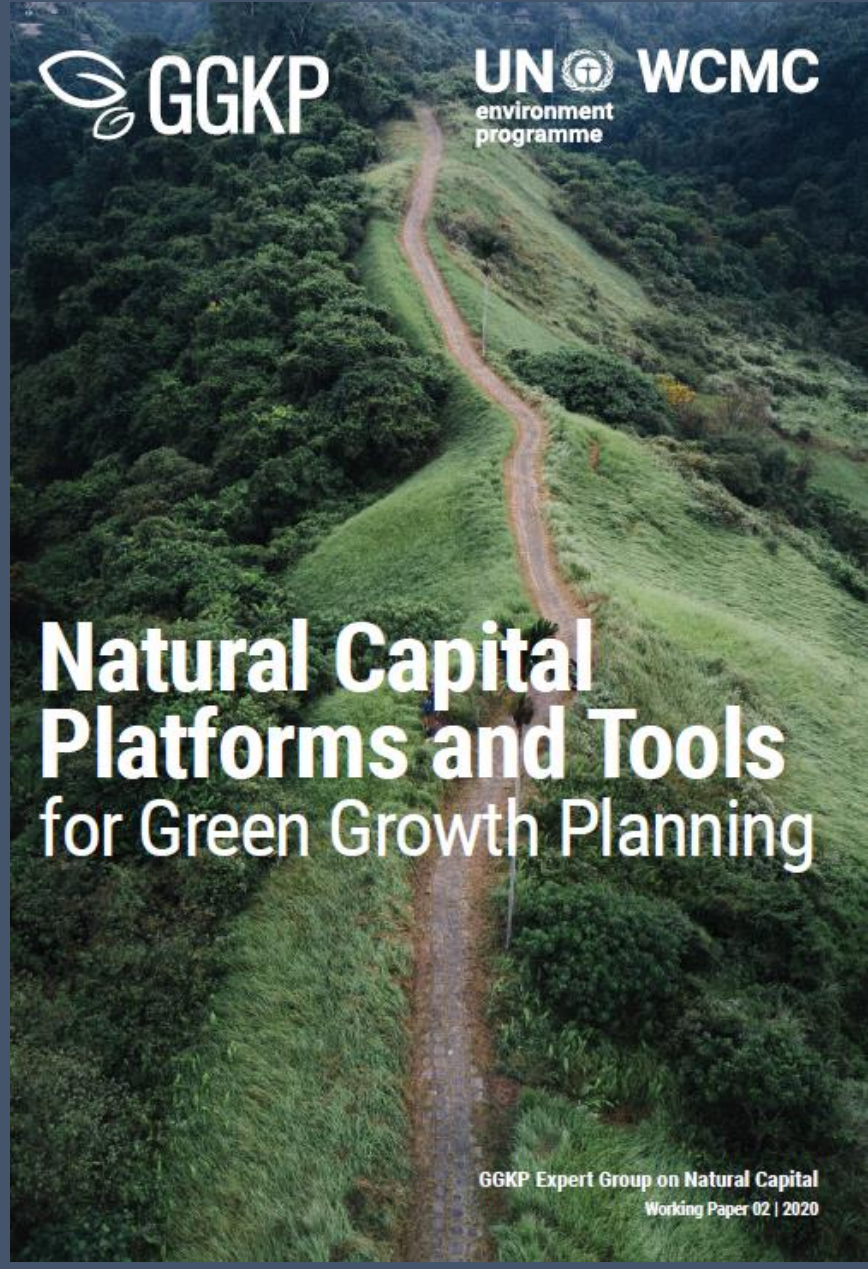
SEEA requires the integration of substantial and diverse data. These include geospatial and other data not originally designed for statistical purposes, whose use is necessary for spatial modeling, which has proven challenging for some National Statistics Offices (NSOs) to implement. A variety of ecosystem

The start of a shared vision?

SEEA accounts & related indicators will be:

1. rapidly recompilable as new science emerges,
2. quickly produced to show the most recent trends as new annual data become available, with
3. robust international comparisons possible from common global data, while country-specific customization is still easily done.

This vision moves high-quality, meaningful information from scientists into the hands of decision makers, the public, and the media as quickly as possible.



GEO EARTH OBSERVATIONS FOR ECOSYSTEM ACCOUNTING

Interoperability requires common goals & standards

- Must go beyond current state of the practice for data & code repositories, APIs, statistical metadata (consistency & machine readability are critical)
- Heiler 1995:
  - Syntactic interoperability: use of compatible data formats and communication protocols. Low bar, more limited advantages.
  - Semantic interoperability: data transfers where a receiving system can properly understand the meaning of exchanged data, reusing it appropriately. Higher bar, greater potential for automation & data/model reuse.

Roles of key stakeholders

- Data providers (NSOs, science agencies, academic scientists): agree on & provide data using common formats & hosting protocols (e.g., SDMX)
- Modelers (science agencies, academic scientists): use modeling practices that will make models more easily linked & documented (more modular, less monolithic); use community consensus semantics
- NSOs & other institutions (NSOs, space/mapping agencies, GEO initiatives, large academic collaborations): maintain interoperable data & models over the long term
- How can we, as a community, envision incentives that will move us beyond the status quo?

Key building blocks for interoperability

1. SEMANTICS: a flexible, shareable, easy-to-learn **language** to describe scientific observations.

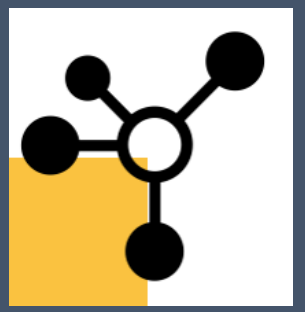
Developed by experts in collaboration with disciplinary scientists – typical scientist/NSO does *not* build these.

Use to accurately describe data & model elements in a consistent, machine-readable way.

2. OPEN, LINKABLE DATA: enabling access & publishing of semantically annotated data.

Put data on the web in machine-accessible formats.

Best practices already exist: no more PDFs of model parameters or zip files of spatial data.



3. OPEN, LINKABLE MODELS: open, accurate, "Wikipedia-like" sharing and linking of models.

Code models in a modular style that facilitates reuse (vs. monoliths).

Build documentation into code for automated reporting.

Specify appropriate conditions for safe reuse of your models.

Moving forward

- We'd like to start working with all stakeholders to address interoperability, in an **inclusive** way.
- Feedback & **ideas** very welcome!

ARIES for SEEA: Next steps

1. Country pilots (India, South Africa, Malaysia, Mozambique, Philippines, Senegal)
2. Training the trainers (UNSD)
3. Virtual regional workshops
4. Feasibility demo for SEEA CF (sdmx integration & sample outputs)
5. Various improvements to interface & models (#1: custom data upload from web Explorer)
6. Adding more models to the Explorer (nature-based tourism, water provisioning)
7. Further outreach to data/modeling community (GEO & others)