Review of data, analytical systems and code for calculating SDG Indicator 15.3.1: Proportion of land that is degraded over total land area. Technical Note.

Version: 1.0
Date: July 2020

LAND DEGRADATION NEUTRALITY INITIATIVE

Acknowledgements
The authors thank those who were able to input into the consultation process for their quick turn around.

Lead Author
Alastair Graham (Geoger)

Contributors and Reviewers
Neil Sims (CSIRO)
Alex Zvoleff (Conservation International)
Sara Minelli (UNCCD)
Antje Hecheltjen (GIZ)

Barron Joseph Orr (UNCCD)
Douglas Cripe (GEO Secretariat)

Suggested citation
GEO-LDN Initiative (2020). Review of data, analytical systems and code for calculating SDG Indicator 15.3.1: Proportion of land that is degraded over total land area. Technical Note. Group on Earth Observation Land Degradation Neutrality (GEO-LDN) Initiative, Geneva, Switzerland.

Table of Contents

Table of Contents 2

Acronyms 4

Executive Summary 7

Introduction 8

Consultation 9
- Methodology 9
- Summary of survey findings 10

Existing analytical tools relevant to SDG 15.3.1 11

Systems for calculating indicators for land-related SDGs 13

Guidance on the development of a Collaborative Data Analytics Platform 15
- What is a Federated system? 15
- Advantages of collaborative frameworks and models 15
  - Autonomy 16
  - Equality 16
  - Access control 16
  - Traceability 16
- Challenges of collaborative frameworks and models 16
- Possible interfaces 17
- Potential existing technical solutions 18

Emerging datasets and code relevant to SDG 15.3.1 20
- Datasets 20
- Code 21

Discussion 21
- Provenance 21
- Licencing 22
- Trust and Approval 22
- Developments 22
- Gaps 23
- Data hosting 24
- Data sharing 24
- Reporting 25

Conclusions 25

Appendix A: Verbatim Survey Responses 27
- Table A1: What comments do you have regarding the possible creation of a federated analytic platform for SDG 15.3.1 data analysis? 27
- Table A2: What SDG specific data processing platforms are you developing or aware of? 30

Table A3: Are you aware of any federated analytic platforms related to the other SDGs? 33
Table A4: What issues around hosting, data security and processing do you foresee with such systems? 33
Table A5: What SDG 15.3.1 specific analytical tools are you developing or aware of, and how will they add to the monitoring and reporting process? These might be GIS tools, Jupyter notebooks, dedicated software, Earth Engine javascript etc. 35
Table A6: What new datasets and/or analytical code are you aware of specifically related to the three sub-indicators? 37
Table A7: What analytical code are you aware of that might be being developed for other SDGs that could be forked and altered to be made suitable for SDG 15.3.1 monitoring and reporting? 38
Table A8: What other comments do you have? 39

Acronyms

Acronym	Meaning
AI	Artificial Intelligence
API	Application Programming Interface
AWS	Amazon Web Services
BDC	Brazil Data Cube
CCI	Climate Change Initiative
CEOS	Committee on Earth Observation Satellites
CI	Conservation International
CLUE	Conversion of Land Use and its Effects
CODE	Center for Open Data Enterprise
COE	Convergence of Evidence
CSIRO	Commonwealth Scientific and Industrial Research Organisation
CV4GC	Computer Vision for Global Challenges
DCoD	Data Cube on Demand
DIAS	Data and Information Access Services
DQS	Data Quality Standards
EO	Earth Observation
EO4SDG	Earth Observations for Sustainable Development Goals
ESA	European Space Agency
EU	European Union
EUMETSAT	European Organisation for the Exploitation of Meteorological Satellites
EVI	Enhanced Vegetation Index
FIS4SDG	Federated Information System for Sustainable Development Goals
FLINT	Full Lands Integration Tool
FORE-SCE	FOREcasting SCEnarios of land-use change
GEE	Google Earth Engine
GEO	Group on Earth Observations

GEOSS	Global Earth Observation System of Systems
GHG	Greenhouse Gasses
GIS	Geographic Information System
GIT	Geospatial Information Technology
GIZ	Deutsche Gesellschaft für Internationale Zusammenarbeit
GPG	Good Practice Guidance
GPU	Graphics Processing Unit
GRID	Global Resource Information Database
HTML	Hypertext Markup Language
JRC	Joint Research Centre
LDN	Land Degradation Neutrality
LULC	Land Use / Land Cover
MDL4EO	Machine and Deep Learning for Earth Observation
MFFESD	Ministry of Forestry, Fisheries, the Environment, and Sustainable Development
MOLES	Multi-Objective Local Environmental Simulator
MPL	Mozilla Public Licence
NASA	National Aeronautics and Space Administration
ODC	Open Data Cube
OECD	Organisation for Economic Co-operation and Development
OGC	Open Geospatial Consortium
ONS	Office for National Statistics
PLUP	Participatory Land Use Planning
RCMRD	Regional Center for Mapping of Resources for Development
REST	Representational state transfer
RLMUA	Rwanda Land Management and Use Authority
RUSLE	Revised Universal Soil Loss Equation
SCO	Science Coordination Office
SDC	Swiss Data Cube
SDG	Sustainable Development Goals
SEN4LDN	Sentinel for Land Degradation Neutrality

SLM	Sustainable Land Management
SNAP	Sentinel Application Platform
SOC	Soil Organic Carbon
STAC	Spatio Temporal Asset Catalogue
TEP	Thematic Exploitation Platform
UNCCD	United Nations Convention to Combat Desertification
UNDP	United Nations Development Programme
UNEP	United Nations Environment Programme
USA	United States of America
VIIRS	Visible Infrared Imaging Radiometer Suite
VISAT	Visualisation and Analysis Tool
WAD	World Atlas of Desertification
WCS	Web Coverage Service
WG	Working Group
WOCAT	World Overview of Conservation Approaches and Technologies
WPS	Web Processing Service
WSF	World Settlement Footprint
WUE	Water Use Efficiency

Executive Summary

This Technical Note provides a review of i) existing analytical tools, ii) emerging data and analytical codes, iii) existing systems supporting the calculation of indicators used in planning and implementation of relevant land related SDGs, and iv) guidance on the potential development methods for a collaborative and modular data analytics platform.

It is important that reporting countries are confident in using a secure and trusted environment to generate the statistics that they need to report into UNCCD, the custodian agency for SDG Indicator 15.3.1, proportion on land that is degraded over total land area. The UNCCD provides countries with a secure environment in which to generate and report LDN statistics. The GEO-LDN Initiative data analytics platform will be a separate platform, distinct from the security surrounding the statistical reporting system.This review is designed to ensure that effort is not duplicated.

A small consultation process has been completed, consisting of an emailed survey of seven questions sent to 13 recipients, chosen following discussions with WG3 leaders and an internet search for relevant affiliations related to SDGs and LDN. A further two land use data practitioners (from Rwanda and Ethiopia) were approached to ask about their requirements. Information from the surveys and from a wider review process has resulted in a number of existing datasets, programs, platforms and initiatives being identified, all of which might have an input into the development of a GEO LDN system of systems.

A series of gaps have been identified that require further investigation. These are:
- Issues around data hosting
- Methods for data discovery and data suitability checking
- Trust and security of an implemented system
- Options for cloud implementations versus local implementations

Further to this, a series of “nice to have” options for a LDN processing system have been suggested. These are:
- The inclusion of sandbox tools
- Some form of upfront cost assessment
- A method for the management of processing tools

Suggestions have been presented on how such a platform might be designed and implemented, although the specifics of design and implementation is outside of this overview and review.

Introduction

This Technical Note[1] provides a review of i) existing analytical tools, ii) emerging data and analytical codes, iii) existing systems supporting the calculation of indicators used in planning and implementation of relevant land related SDGs, and iv) guidance on the potential development methods for a collaborative and modular data analytics platform.

This review is driven by a number of different components. It is important that reporting countries are confident in using a secure and trusted environment to generate the statistics that they need to report into UNCCD, the custodian agency for SDG Indicator 15.3.1, proportion on land that is degraded over total land area. It is important to balance the global data perspective, by using comparable and standardised datasets, with that of national data sovereignty and ownership and the demand to produce and report data that are relevant to national and sub-national contexts, including for implementation and support of planning and decision making processes. Part of the solution includes data quality standards (DQS) and good practice guidance (GPG), developed in conjunction with an analytics platform. The UNCCD provides countries with a secure environment in which to generate and report LDN statistics. The GEO-LDN Initiative data analytics platform will be a separate platform, distinct from the security surrounding the statistical reporting system.This review is designed to ensure that effort is not duplicated, but rather that efforts are reused and improved on.

From the point of view of the custodian, it is important to achieve reproducibility of data processing and comparability of results across all reporting entities. For this to be completed, country specific data need to be tested against common data standards whilst users must also have access to a range of trusted global datasets, that can be used for reporting in the absence of, or to complement, country specific data. A further part of the process is to use open, accessible methods and software throughout the processing chain, so that all components of the reporting process can be assessed. With these factors in mind, this review assesses where gaps exist in the potential creation of a collaborative and modular data analytics system, and the data which might form part of such a system.

Secondary requirements for the formation of a collaborative system might include issues around software sustainability[2] and collaborative development working techniques. Software sustainability pertains to the methods employed to ensure that copies of working software are preserved, and that different versions are maintained in a curated way. Specific software sustainability methods involve, amongst others, technical preservation, migration to modern systems and systematised deprecation, and these should be considered at the start of the WG3 process. Free and open code aids the concept of software sustainability by being visible to all. Collaborative development is well known in the software and science sectors, and involves the acceptance of inputs from multiple individuals and agencies. It is strongly recommended that any data analytics and modelling system developed in WG3 considers these aspects, not just to allow inputs from the developers and data providers but through a feedback loop that involves the users and analysts.

[1] This technical note has been created and is provided for the GEO LDN Initiative Working Group 3 team.
[2] https://www.software.ac.uk/

Consultation

Methodology

The first step in this reporting and assessment work involved a small consultation process. This consisted of an emailed survey of seven questions sent to 13 recipients, chosen following discussions with WG3 leaders and an internet search for relevant affiliations related to SDGs and LDN. The people contacted are shown in Table 1 along with the relevant technology they are linked to and whether they responded to the survey. The survey turnaround was short, with those approached asked to provide their opinions within seven days, and the author thanks those who responded with their inputs. Those listed in Table 1 were contacted based either on recommendation from the WG3 team leaders, or through discovery on their work via internet searches.

Table 1: Survey respondents and their affiliations[3]

Survey respondent	Technology/project represented
Gregory Giuliani[4]	Swiss Data Cube Ported Trends.Earth LDN code to their data cube
Xiaosong Li	Chinese Academy of Sciences Pushing forward Chinese response on LDN
Marc Paganini	SEN4LDN Pushing forward the Sentinel for LDN project
Michael Cherlet	JRC / WAD Working on LDN tools for the World Atlas of Desertification
Brian Killough	Open Data Cube Has been implementing LDN code in Python notebooks on data cubes
Emil Cherrington[5]	SERVIR A global network for environmental monitoring
Rob Waterworth	FLINT Involved in developing the FLINT open source platform
Philip Frost	iMMAP Have been implementing some of the SDG codes onto their data cube
Mariano Gonzalez-Roglich	Tools4LDN Improving tools and methods for LDN

[3] Three additional organisations were approached for input, representing openSDG, EO4SDG and FIS4SDG.
[4] Pierre Lacroix and Pascal Peduzz from MapX stated that they had spoken with Gregory Giuliani and were confident that Gregory's response would reflect their views.
[5] Emil would like to recognise other SERVIR Science Coordination Office (SCO) colleagues who provided input. Africa Flores serves as the Land Cover, Land Use, and Ecosystems Lead at the SERVIR SCO (and Regional Science Coordination Lead for Amazonia), Emily Adams is the Regional Science Coordination Lead for Eastern & Southern Africa - and provided much of the input regarding RCMRD's LDN-related work - Francisco Delgado is the lead for the SERVIR SCO GIT team, Dr. Ashutosh Limaye is the SERVIR SCO Chief Scientist, and Eric Anderson is the Associate Chief Scientist. What is provided summarises some of their discussions.

Summary of survey findings

This section provides a summary of the amalgamated responses to the survey questions. The verbatim responses are presented in Appendix A at the end of this document.

In terms of the generation of a data analytics system, the prevailing view from those approached is that this would be a laudable end goal with benefits to those who develop and use it. It is beyond the scope of this document to request user input, but following the concept of "design with the user" an overview on what is being used at a national level would be an interesting addition. Some of these national tools could also be considered for inclusion in an LDN system of systems.

The specification of such a system (including how it is managed as it develops) is critical. Issues surrounding data discovery, technical systems interoperability, data supply, data processing, data security, the supply, management and quality assessment of training and validation data were all raised by the survey respondents as topics to consider.

Making sure that suitable data, which adheres to the data quality standards defined as part of the GEO LDN Initiative's WG2 work[6], is used in such a system relies on there being a relevant *data discovery tool*. A *quality assessment tool* based on the WG2 data quality standards decision trees could also be a useful addition, enabling users to assess their national and sub-national datasets as they are brought into the analytical system. Linked to the issues surrounding data quality, are those relating to understanding the *quality of any training or validation data*. This is critical in any Earth observation application, but particularly so if the analyst chooses to undertake machine or deep learning techniques. Being able to assess and communicate the precision and accuracy of the training and validation data, as well as the processing result, would be a key benefit to understanding the efficacy of the generated product.

Potentially related to this topic is the link between local and global data. Often, field data are collected locally for purposes that are related to, but not part of, the LDN proces. These could befor carbon certificates, project monitoring, environmental research, training algorithms, etc. A platform that was able to link local, national and global data could provide an important development.

In terms of data security, the respondents have suggested that data remains with the data provider or curator, and that there needs to be trust in the method with which it is transmitted to each hub or user. A related item not mentioned in the responses surrounds the specifics of data protection and privacy. This would be a key consideration in EU countries and one that would potentially affect many other countries. The national politics surrounding data protection and security will also need to be considered, as there may be teams within the LDN process who are unable to collaborate fully.

Other suggestions that arose from this survey were that there should be an option to include and use *sandbox tools* (i.e. an area for trialling new data, algorithms or methods, and comparing the outputs with the standard methods used for LDN reporting), and that repeatability of results should be built into the system (e.g. for instance, being able to create, download and share a data processing recipe).

---

[6] GEO-LDN Initiative (2020) Minimum data quality standards and decision trees for SDG Indicator 15.3.1: Proportion of land that is degraded over total land area. Technical Note, Group on Earth Observation Land Degradation Neutrality (GEO-LDN) Initiative, Geneva, Switzerland.

In terms of existing data processing systems, the following were mentioned:

- the Swiss Data Cube (https://www.swissdatacube.org/)
- Google Earth Engine (https://earthengine.google.com/)
- GEOSS and the VirtLab[7] (https://www.earthobservations.org/geoss.php)
- CASEarth (http://www.casearth.com/index_En.html)
- EuroDataCube (https://www.eurodatacube.com/)
- the DIAS systems (https://www.copernicus.eu/en/access-data/dias), and
- the Thematic Exploitation Platforms (TEP) (https://eo4society.esa.int/thematic-exploitation-platforms-overview/)

It is recognised that not all of the solutions listed are analytical platforms in and of themselves, but contain some form of data hosting or the capability to host an analytics system.

Following on from these suggested platforms, the issue of data hosting and security was raised. It was stated that some countries would have difficulty in using (and submitting data to or through) commercial services such as AWS and the Google Cloud due to privacy concerns and local laws. The data supply needs to be neutral, with data consistency across the various hubs of a collaborative system: a potential way forward could be for data providers to supply data through OGC web services, and to allow for local data sources to be included in the analytical system. It was also suggested that any platform developed should include hubs that are installable across different hosting levels for maximum flexibility (e.g. local system, cloud services etc.). Gaining some form of collaboration at the build stage, with users and other developers, will increase trust and uptake of the resulting platform and it was suggested that it might be beneficial to have the code managed under a trusted and relevant umbrella organisation.

In terms of tools mentioned by the respondents, the Swiss Data Cube has ported some of the Trends.Earth code into its system[8], a SDG 15.3.1 reporting tool is mentioned in regards to Google Earth Engine, Jupyter and Colaboratory notebooks are mentioned, as is the VISAT data exploration portal which is part of the urban TEP.

It was reported that a 30m global data product has been produced in China, covering the time period 2000 - 2018. In terms of Sentinel data, the SEN4LDN project is aiming to produce Sentinel based products of relevance to LDN reporting, and a Sentinel-1 data cube is also in the planning stages.

Issues around choosing the correct software and data licences (open source code and open data should ideally be licenced for clarity, using one of the *many* open licences) were raised, as was the concept of having a flexible user interface that changes depending on the type of user that is logged into the system.

Existing analytical tools relevant to SDG 15.3.1

Two recent publications (Schrodt *et al.*[9], 2020; Szantoi *et al.*[10], 2020), although not directly targeting SDG 15.3.1, provides a useful overview of methods and data available for land based ecological monitoring. With these publications in mind, figure 1 shows how the data analytics platform will fit in the GEO LDN Initiative, and wider context, as well as how different tools, ideas and programmes will help

[7] https://www.mdpi.com/2072-4292/12/11/1795/htm
[8] https://doi.org/10.1080/20964471.2020.1711633
[9] https://link.springer.com/chapter/10.1007/978-3-030-33157-3_17
[10] https://doi.org/10.1016/j.envsci.2020.04.005

guide the platform development. As is evident, a lot of initiatives are underway under the LDN banner, but there are opportunities for other projects to interface with parts of the process; specifically, the data analytics platform.

Figure 1: The wider context of a collaborative, modular data analytics platform.

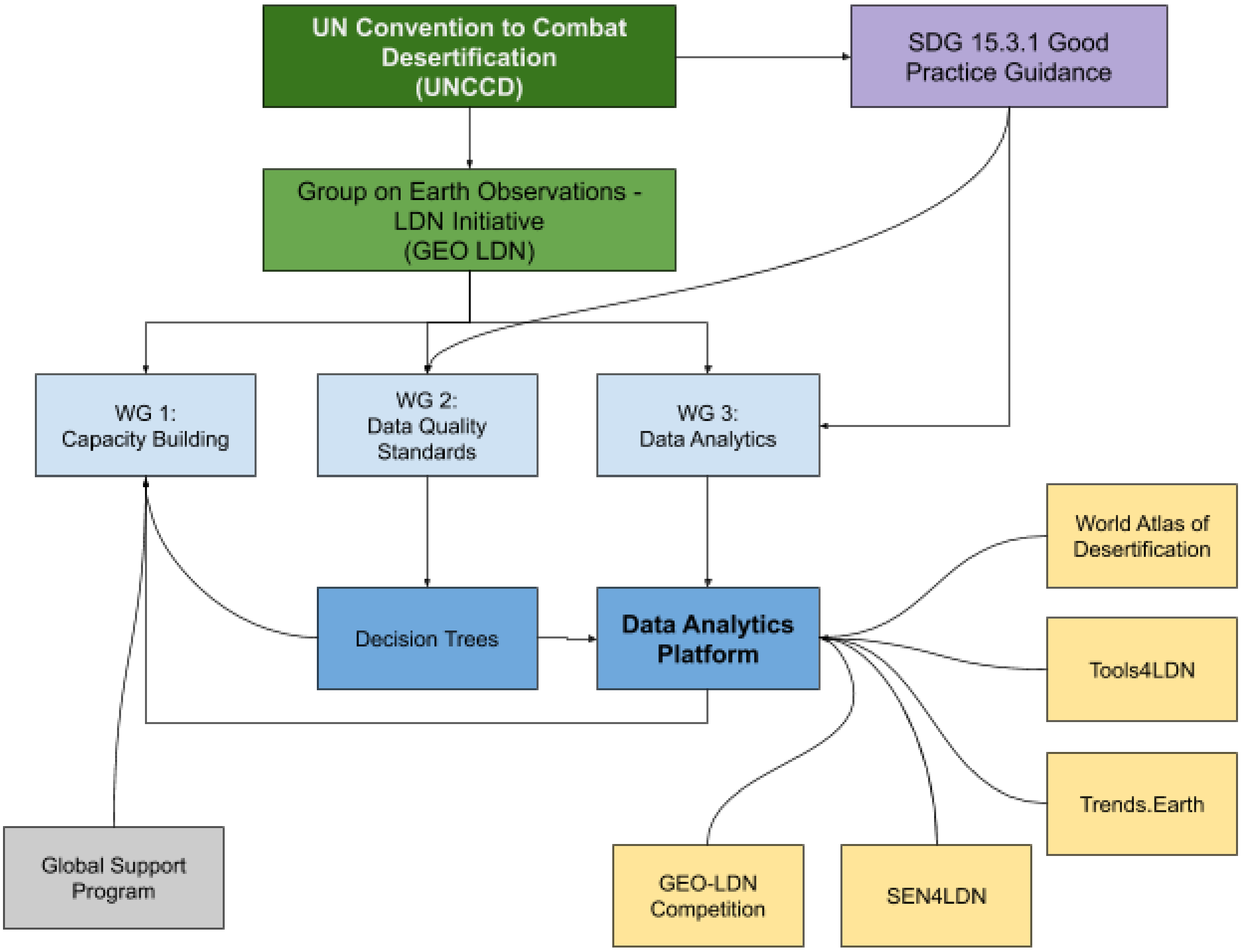


Table 2 lists some of the existing tools and projects that are currently available for potentially calculating SDG 15.3.1 sub-indicators.

Table 2: Existing tools and projects.

Tool / Project	Notes
trends.earth	Trends.earth is implemented as a QGIS plugin that links with Google Earth Engine and is the current de facto SDG 15.3.1 analytical tool.
Tools4LDN	This project is looking to integrate the WOCAT SLM database, LandPKS mobile application and Trends.Earth QGIS plugin with additional socio-economic and drought indicators. This is a critical new project with the potential to have a real impact on how LDN is monitored and reported.
EO4SDGs	The EO4SDG project presents information about the potential of Earth observations and geospatial data for the Sustainable Development Goals. It presents case studies from other projects that are directly working on

	solving the SDG problems[11].
SEPAL	SEPAL is part of the Open Foris suite of tools. It is designed to process satellite data and create products suitable for local needs. To do this SEPAL uses cloud-based services such as Google Earth Engine. It also uses open source software including Orfeo Toolbox, Gdal, R, R Studio Server, R Shiny Server, SNAP Toolkit, and the OpenForis Geospatial Toolkit.
Collect Earth	Collect Earth is a cross platform tool that allows data collection through Google Earth, Bing Maps and Google Earth Engine, and is also part of Open Foris. Users can interact with the customisable tool to analyse satellite imagery. Collect Earth Online is a webenabled version of the tool which is available through Digital Earth Africa[12] for data validation.
Google Earth Engine plugin for QGIS	Although trends.earth already interfaces with GEE, this dedicated plugin has been created to enable direct coding in javascript in GEE via a QGIS interface.
MapX	MapX has been developed by UN Environment, the World Bank and the Global Resource Information Database (GRID-Geneva) to use cloud computing for the assessment and management of natural resources. The platform attempts to bring together data held by different providers to allow better decision making.

Systems for calculating indicators for land-related SDGs

In general, land change models are developed to describe and explain changes in the dynamics of land use and land-cover[13]. They can be used to explain collected time series of data, as well as projecting likely changes into the future. There are different methods that such models can employ, including machine learning and statistical models, cellular models, economic models, agent-based models and hybrid models (which combine many of the former model types).

A paper by De Rosa, Knudsen and Hermansen (2016)[14] assessed the different methodological approaches of land use change models. Six models were used in the review and were compared according to fifteen criteria covering: modeling framework, impact categories assessed and model transparency. The results demonstrated that progress was being made in the development of Economic General Equilibrium Models. It was concluded that LULC models can still be improved, especially when considering the "precision of data, identification of marginal land and inclusion of a broader range of impact categories". It was stated in the review that existing models (at the time of the paper's research) focus on GHG emission-related impacts and that other environmental and socio-economic impacts need to be considered.

To better understand what methods are currently being used to assess land use change and how such methods might be used to understand future planning of land resources, linked to climate and land based changes, two practitioners were identified and approached[15].

[11] The data4SDGs project is taking a wider view and investigating the potential of a variety of data sources.
[12] https://www.digitalearthafrica.org/index.php/news/collaborative-tool-removes-barriers-data-validation-africa
[13] https://en.wikipedia.org/wiki/Land_change_modeling
[14] https://doi.org/10.1016/j.jclepro.2015.11.097
[15] These practitioners' contact details were made available by GIZ.

The first practitioner was Genanaw Alemu, a GIS, Remote Sensing Advisor in the Participatory Land Use Planning (PLUP) Project run out of Addis Ababa, Ethiopia. He reported that high resolution satellite imagery from commercial suppliers was being used to extract existing land use and land cover (LULC) information and then carry out land use planning at a larger scale. Data from open sources, such as SRTM DEM, was used to provide additional information such as elevation and slope for target areas. The project depends on open source software such as QGIS, although it was stated that many spatial data specialists in Ethiopia are more familiar with ESRI and Hexagon products than QGIS. Currently, no modelling software is being used in the production of land use planning outputs.

The second practitioner approached was Alexis Rutagengwa, the Head of the Land Use Planning, Mapping and Surveying Department at the Rwanda Land Management and Use Authority (RLMUA).
He reported that geodata is widely used in the RLMUA, especially for land administration processes and in land use planning and management work. The main software being used to undertake work related to land administration and land use planning is ArcGIS (both ArcGIS 10.7 and ArcGIS pro). These data and software have been used to update the national basemap (comprising an updated national geodatabase and high resolution satellite imagery, resulting in an updated land cover dataset for 2019 and production of topographic map sheets at 1:50,000 scale). The National Land Use and Development Master Plan has also been revised to guide land use development in Rwanda for the period 2020-2050. Further, the spatial cadastral database, which is hosted in the Land Administration Information System, is being used to understand the current land use for all land parcels. The same system is being used to display and publish the planned land use information from different master plans. A land use decision support tool has also been developed to be utilised across different governmental and land planning institutions. The use of land use modelling software is limited but a collaboration with the Rwanda Water and Forestry Authority has produced an erosion risk map using RUSLE[16] factors which captures the "high to extremely high" erosion areas across the country. A 2019 National Forest Cover map has also been created using high resolution satellite imagery (50cm) over the past year, the Ministry in charge of disaster management has conducted a drought susceptibility mapping exercise. All these data are being used in land use planning.

Despite no specific land use modelling being used by these practitioners, there are different models available that could be used in some way to inform land use change planning and land degradation.

Although not directly linked to SGD 15.3.1, the OECD Multi-Objective Local Environmental Simulator (MOLES)[17], is an urban Computable General Equilibrium model which links urban land use, mobility patterns, projected emissions and potential environmental impacts. This model has been used to assess different urban development scenarios through to 2050, and includes a geospatial data component. Further investigation of this model and the components used in it could be useful, with the potential for collaboration between teams looking at different (but related) aspects of land use and land cover change.

The USGS have been researching local and global drivers of land-use change to better understand how different scenarios might impact on landscapes. The FOREcasting SCEnarios of Land-use Change (FORE-SCE) modeling framework[18] has been created to allow spatially explicit projections of future land-use and land-cover change for parts of the continental United States. FORE-SCE uses different

[16] https://efotg.sc.egov.usda.gov/references/public/WA/Revised_Universal_Soil_Loss_Equation_(RUSLE).htm
[17] https://www.oecd-ilibrary.org/environment/multi-objective-local-environmental-simulator-moles-1-0_151cf08a-en and https://ec.europa.eu/newsroom/know4pol/document.cfm?doc_id=64007
[18] https://www.usgs.gov/land-resources/eros/lulc

scenario datasets and spatial allocation methods to create different spatially explicit maps of land use change. The model has been developed internally for use by USGS researchers and the software is not openly available, although USGS have stated that they are willing to discuss potential opportunities for collaborative research applications.

Finally, the Conversion of Land Use and its Effects (CLUE)[19] models "are among the most frequently used land use models globally". The application of the CLUE models can be geographically disparate, from local and regional, to global. The base model is both flexible and generic and is distributed as freeware (rather than being open source).

Guidance on the development of a Collaborative Data Analytics Platform

What is a Federated system?

One of the ideas for a collaborative, modular system of systems has been to investigate the utility and relevance of implementing a federated system. It must be stressed that this would only be one of a number of possible solutions.

The concepts of a distributed system and a federated system are commonly confused. A distributed system is typically one where deployment is made over a geographical area. For example, in the case of retail, an organisation might operate a number of stores that are geographically distributed away from a head office. A distributed IT system should be able to manage, deploy and upgrade any location with no organisational overhead as well as having the capacity to provide high availability without the need for expensive infrastructure[20]. A well constructed distributed system will also have the ability to send processes across a network to perform data processing at a location that is most optimal e.g. next to the data repository.

A federated system is a subtly different concept: different organisations/softwares are brought together in such a way to create a coherent overall system. Critical technologies for such a system include software designed for identity and process management. It is not imperative that a federated deployment is distributed, but this is often the case in large operational systems as this can lead to multiple optimisations across the various components.

Advantages of collaborative frameworks and models

There are certain advantages in the use of collaborative (and federated) frameworks and models. A selection of these are listed in this section. First, it is important to define what the required output is for such frameworks - one definition can be that users are looking for the "coordinated interchange of information among independent data hubs that allows for data sharing while maintaining control over

[19] http://www.ivm.vu.nl/en/Organisation/departments/spatial-analysis-decision-support/Clue/index.aspx
[20] https://www.progress.com/blogs/distributed-vs-federated

own data resources”[21]. There are four key concepts to be addressed: autonomy, equality, access and traceability.

Autonomy

When discussing federated systems, autonomy relates to the concept that each hub is free to join or leave the federated system. Linked to this membership of the system is whether the hub is sharing or withdrawing access to previously shared data. It would not be beneficial to the overall system if hubs were able to withdraw data access rights without warning; but conversely, each hub requires the right to decide on what to share and for how long.

Equality

The success of a federated system is in part due to the equality between each hub. In a true federated system, no hub should have authority over another, even if some hubs are more popular with users than others.

Access control

This concept is closely related to that of autonomy. In general, each hub should be able to decide what content to share, who can access that content, and by what means that content can be accessed. However, any federated system related to the LDN process would likely need to be considered in a slightly different manner. Certain data (such as results) must be shared for the system to meet its goals, although in the case of the results, hubs and users would only be sharing with one trusted entity. However, there is also the potential that regional hubs could be constructed to share regionally relevant datasets (that meet minimum data quality standards) between multiple countries and users.

Traceability

This key aspect of a federated data system relates strongly to the issues of trust and security. The data assets should remain with the providers, but references to them may be passed to and between the other hubs. Traceability of what data are used, how, and from what source should be automatically captured in the metadata of the analytical process for each country.

Challenges of collaborative frameworks and models

Some major challenges exist around creating and maintaining collaborative frameworks and models.

The first of these relates to the hosting of the system. There are three components to this: technically, the hosting location needs to be able to accommodate all of the requirements of the system; financially, the costs of creating and operating the system need to be manageable; and politically, the hosting location needs to adequately address the issue of trust if the system is to drive uptake by different user groups. With so many countries participating in the LDN process, being able to address all of the concerns regarding hosting location will be key. For instance, some countries may not be comfortable using cloud providers such as AWS and Google Cloud, whilst others may require certain geographies to be restricted. Add more

[21] http://ggim.un.org/meetings/2019/Deqing/documents/S3-FIS4SDGs_L%20G%20Morales.pdf

The second challenge comes when attempting to maintain version control between different hubs within a system. There are two components to this: version control of the underlying software, and versioning of the data. Version control is important for allowing users to know where in the development chain their hub sits. Being able to upgrade as and when required by the user is also an important consideration for such a system - although being able to enforce upgrades[22] at specific times can also be critical for the developers and data providers.

A third challenge is that of politics: in this case both national politics but also that of individuals and groups when it comes to the issue of data and code sharing. Being willingly able to provide open data and open code and models that may be relevant to the wider overall project will help develop an ecosystem of users that share best practice and help drive forward developments. This co-ownership of the system is an important aspect to aim for, but will likely take much effort to attain.

Possible interfaces

A major consideration when creating a system to be used by a large number of people with unknown technical abilities, across different languages and from different cultures is how to design the user interface. Indeed, a complete report could be written regarding user interface design and how these are best presented to maximise user interaction with the software; but this is not the focus of this review. There are, however, a number of different ways in which it is envisaged that users could interact with a data analytics system depending on what task they need to undertake. These include the following:

- Traditional GIS
    - One of the most familiar and standardised interfaces that analysts might use in such a system is that of the traditional GIS. The system, particularly if an open-source GIS software package, is extensible. Whether the data are hosted on the local system, on an on-premises server system or delivered through the cloud makes little difference and is part of the flexibility of a GIS system. One benefit of maintaining a traditional GIS system is that the underlying spatial and mapping components can be leveraged into any analysis offering required to create relevant outputs. The trends.earth plug-in currently sits within the QGIS software framework.
- Python notebooks
    - Jupyter and Collaboratory notebooks allow rapid prototyping of ideas and sharing of results. They are designed for idea and code sharing. They are ideal user interface tools for sandbox applications or for capacity building.
- Online map interface
    - An online portal provides another interface that could be used in any developed system. The functionality of such portals is often limited, but if the ability to be remotely and/or locally hosted was provided, then the power of such an interface could be increased.
- Business intelligence front end
    - A popular interface for non-technical users is the business intelligence dashboard[23]. A business intelligence front end can be hosted remotely or on the analysis system itself, and if sufficiently flexible can be customised to show the relevant information for the specific user viewing it.

[22] An example of this includes the enforced upgrades to the Zoom platform during the global Covid-19 pandemic, although users were also encouraged to make incremental upgrades as releases were made available.
[23] https://logz.io/blog/business-intelligence-tools/ & https://www.metricfire.com/blog/top-8-open-source-dashboards

- Plots
    - A simple, static data portal could be constructed as a user interface, composed of javascript/Python based graphical plots and HTML. This is a limited but easy to understand interface type.
- Data delivery
    - Being able to export the required information on each of the sub-indicators, or from land use planning models, in a standardised and simple format (such as comma separated values [csv], or as an API) could create a lot of flexibility in the system.

Potential existing technical solutions

This section attempts to outline the existing technical solutions that are currently available to be placed into a collaborative and modular system of systems. Each of the following technologies (Table 3) has its own benefits and drawbacks but it is hoped that this overview provides insight into which technologies and softwares might be of actual practical benefit to the LDN process. An academic study[24] into some of the technologies listed in this section has been completed by researchers working on the Brazil Data Cube project, and this may present a more detailed assessment than this Technical Note.

Table 3: Existing platforms and tools which could be of benefit to developing an LDN data analytics system. All options suggested here would require dedicated further investigation prior to inclusion in an LDN system[25]
(key: 🖥️ = platform, 💿 = data repository, 🔨 = toolset).

Name	Function	Description
Earth AI	🖥️ 💿 🔨	EarthAI is a commercially available integrated imagery analysis platform which exists to simplify data discovery, access, and processing specifically related to Earth observation data. The system uses a proprietary Python notebook (EarthAI Notebook) which is hosted in the platform's JupyterLab environment. A dedicated open source toolkit for geospatial cloud computing (RasterFrames) allows analytics to be scaled and customised. The system is GPU enabled, with no restrictions on use and provides API access to petabytes of data.
Cascadia	🖥️ 💿 🔨	The Cascadia Partner Forum is constructing a regional analysis tool for the Cascade Mountains (bordering Washington State, USA and British Columbia province, Canada) using Google Earth Engine. The tool allows "partners to monitor trends over time, project future trends and risks, inform spatial impacts of local projects to the region scale, and coordinate efforts across boundaries. The tool is built to support the presentation of spatial information for identified regionally valued species, ecosystems, and/or ecosystem processes".
Microsoft Planetary Computer	🖥️ 💿 🔨	Primarily designed to answer biodiversity and health questions, the Planetary Computer will aggregate global environmental data, including satellite imagery, and make this available alongside machine learning

[24] https://doi.org/10.3390/rs12081253
[25] Inclusion of a technology in this table should not be taken as a recommendation for use.

		tools and ESRI mapping products. It is a continuation of Microsoft's AI for Earth program.
FIS4SDGs		This aims to be "a platform for the sharing of national and global statistical and geospatial data for the 2030 Agenda" that will create a federated network of national hubs. This project is very relevant to the LDN platform requirements, and should be approached regarding co-development opportunities.
EuroDataCube		The Euro Data Cube is an amalgamation of several service components, such as a data archive of satellite imagery and other geospatial datasets, an analysis platform, a distribution hub and a marketplace for geospatial applications. A related project is xcube which provides a toolset for xarray specifically for manipulating the types of data inside Euro Data Cube.
Google Earth Engine		Google Earth Engine combines a huge amount of satellite imagery and other geospatial datasets with a huge capacity for computational processing. Earth Engine also provides APIs and tools such as a simple to use Javascript coding interface to enable the analysis of large datasets.
DIAS		There are five DIAS platforms (4 dedicated to land applications) which host a variety of satellite and environmental data. These systems provide a hosted data and compute environment and allow developers to build tools within their environment.
Actinia		Actinia is an open source REST API[26], effectively enabling the GRASS GIS software to be packaged and accessed in cloud environments. It is specifically designed to process (time series of) satellite images and other geospatial data.
FLINT		FLINT is an open source data integration tool, maintained under the Linux Foundation's moja.global project, used primarily to calculate and track greenhouse gas emissions from changes in land use. A key differentiator for this software platform is the flexibility it has in how it is installed and made available, and in the fact that it can run future predictions. A commercially hosted and developed version called FLINTpro is also available.
Data 61 tools		Data 61 works with CSIRO scientists to create practical, and innovative, web enabled tools. They have created a number of data and processing portals, based around a number of environmental issues.
Open Data Cube and DCoD		The ODC is a geospatial data management and analysis software project designed specifically for use with large amounts of satellite imagery. Its core technologies are Python and PostgreSQL, and it is an increasingly implemented technology. However, installing and configuring the ODC is difficult and requires experience and knowledge. The Data Cube on Demand (DCoD) approach is a new project that is attempting to reduce these blocking issues. DCoD will allow users to specify some basic requirements of the ODC and it will then launch and manage the data infestation for

[26] A how-to is provided here: https://live.osgeo.org/en/quickstart/actinia_quickstart.html

		that system. DCOD has been successfully tested in two sites in Bolivia and The Democratic Republic of Congo.
xcube	🔨	xcube provides a Python toolkit that uses xarray.Dataset implementations to convert Earth observation data sources into self-contained data cubes for further processing and sharing.
OpenEO	🔨	OpenEO is working to create a single unified open API that can be used to connect different coding languages and clients to big Earth observation data on the cloud. This tool could be useful for ensuring data provision is simple to implement.
STAC	🔨	The Spatio-Temporal Asset Catalogue (STAC) is rapidly becoming a standard for data discovery tools. Creating STAC compliant datasets will help tools and users find data quickly and easily. An ecosystem of associated projects is starting to grow around the STAC requirements.
Magda	🔨	Magda is a data catalog system that provides tools to enrich, search, track and prioritise data. These data can be from multiple different sources.
EODataDown	🔨	This is an open source data discovery and download tool which collects open satellite imagery from the Google Cloud. This can then be imported into tools such as open data cubes.

Although not of direct relevance to an LDN system, the BigGIS project[27] could be worth contacting for discussions around how they are implementing their data and analytics platform. Similarly, PanGeo[28], a community and platform built around Python data processing for the geosciences, might have valuable inputs into the best practice surrounding the construction of a large collaborative system.

Emerging datasets and code relevant to SDG 15.3.1

Datasets

A recently written report on datasets from the Tools4LDN team lists a number of datasets of potential interest to the LDN process.

A Medium post[29] from late 2019 lists a number of datasets, and sources of datasets that could be of potential use to the future calculation of SDG 15.3.1. An initial investigation suggests that many of the datasets listed are insufficient in their geographical or temporal coverage, or do not meet the LDN data quality standards. However, this could change for future iterations of some of the datasets. The post also lists frameworks and platforms relevant to environmental monitoring.

Overall, it is not thought that there are any datasets of direct relevance to LDN that members of the GEO LDN Initiative are not currently aware of.

[27] http://biggis-project.eu/biggis-docs/
[28] https://pangeo.io/index.html
[29] https://medium.com/@eirinimalliaraki/data-for-environmental-intelligence-3f2f17cf8f56

Code

In June 2020, Andrea Markos (based in Bolivia) was contacted about work he was involved in relating to SDG 15.3.1. His current work has been partially funded by the 'Foundation for the Conservation of the Chiquitano Forest' with a UNDP grant intended to fund the design of a post-fire restoration plan. Andrea is currently collaborating with Gregory Giuliani (Swiss Data Cube).

He is implementing land productivity change assessments in GEE which involves creating EVI time series and calibrating with the Water Use Efficiency (WUE): the EVI shows the scale of change and the WUE demonstrates the cost, in terms of water use. Now that Sentinel 2 has a 4 year time series Andrea is keen to use this for such work. However, there is a jump in values between the Landsat series and Sentinel 2 series, which raises questions about using both datasets together. Using Sentinel 2 imagery over large areas is computationally intensive but it can be done. At the time of writing this review, the work that is being undertaken by Andrea is under embargo as he writes up an academic paper on the topic, but remaining in contact with him or inviting him into the WG3 process would be beneficial to developing the LDN analytical process.

Some work on SDG 15.3.1 has been undertaken using Ukraine as a case study, as part of the ECOPotential project[30]. The code for this is available on GitHub[31], and is Python based but designed to be run in VirtLab.

Two groups to remain aware of, although they appear to have been less active in 2020, are Machine and Deep Learning for Earth Observation (MDL4EO)[32] and Computer Vision for Global Challenges (CV4GC)[33]. Both of these groups are collaborations of researchers looking to use various subsets of AI to investigate specific issues that could be of use to future LDN developments. For instance, some of the MDL4EO work has looked at methods of data fusion around different satellite sensors, something that has been raised in the WG2 Data Quality Standards work as being of interest to the LDN community.

Discussion

Provenance

As this brief review has highlighted, there is an opportunity for the development of a (potentially federated) collaborative and modular system for LDN analysis. In some respects this also applies to methods that could potentially be used in implementing LDN reporting.

For instance, one technology that could be used in both data transmission and trust, as well as securing the provenance of country-specific reporting is blockchain. Although still in a very early stage of implementation in the geospatial and more specifically the Earth observation sector, blockchain does have a strong potential role here in enabling all reporting parties to understand where data (or reported statistics) have originated and whether or not they have been altered. The blockchain is an open

30 https://doi.org/10.1080/17538947.2019.1610807
31 https://github.com/LeonidShumilo/Vlab1531
32 https://mdl4eo.irstea.fr/
33 https://www.cv4gc.org/

distributed ledger that can record transactions between two parties efficiently and in a verifiable permanent way i.e. the transaction of data to individual countries, and statistics back to the UNCCD could benefit from this technology.

Similarly, an easier to implement solution could be to create hashes with all submitted documents. Many people are familiar with the concept of hash files through the sharing of MD5 Checksum files when downloading software. The MD5 Checksum creates a hash: a unique file specific to the document or file it is protecting. The document or file gets transmitted separately to the MD5 file. The recipient recreates the MD5 file and then checks it is exactly the same as the original hash: if not, then the transmitted file has been corrupted or altered intentionally. This is a method that is part of the blockchain/data security model, but which can also be applied in a simple form[34]. By checking the provider's hash against the receiver's hash, it is possible to know whether any file has been altered since leaving the provider.

Licencing

It is important to apply the correct open source licence to the various software components utilised in any system that is developed. A data licence (e.g. Creative Commons) should also be applied to the open datasets used in such a system.

It was mentioned in the survey responses that FLINT is licensed using the Mozilla Public Licence 2 (MPL2). This "is a copyleft license …[where a user]... must make the source code for any ... changes available under MPL, but ... can combine the MPL software with proprietary code, as long as ... the MPL code …[is kept]... in separate files"[35]. This and other licenses should be fully assessed for their appropriateness for the LDN process.

Trust and Approval

One of the considerations around the creation of a collaborative, modular and potentially federated system and the provision of specific datasets, is how to get approval from users? This is a difficult question to answer, and one that takes time and demonstrates that the system is usable, stable, secure and trusted by others. However, maybe it doesn't need answering at present. The GEO LDN process has a large existing community of generally highly engaged and enthusiastic users: working with WG1 (capacity building) and being open about all aspects of systems development, it may be simple to maintain and grow the existing trust and approval for an updated system. Developing in a non-partisan, flexible manner is encouraged.

Developments

One aspect that would be useful for the wider process of progressing towards meeting the SDGs is for there to be some related way of sharing lessons learnt between different SDGs. Ideally this would utilise an existing platform or technology, would be moderated and would include an API that would allow messaging to be undertaken from within the LDN data analysis platform. This would allow different users from different countries to interact and share best practice or pitfalls, as well as allowing other SDG practitioners to join as their reporting processes became as developed.

[34] https://www.tutorialspoint.com/data_structures_algorithms/hash_data_structure.htm
[35] https://tldrlegal.com/license/mozilla-public-license-2.0-(mpl-2)

A further addition to the system to be considered, is a way for users to add to the system. This could be via the generation and sharing of new datasets, via the creation of new code and/or code snippets, and potentially the suggestion of new methods as investigative and analytical work is undertaken at a national and sub-national level.

Gaps

Table 4 lists the main gaps that have been identified as part of this work. It is suggested that there are four main areas requiring further investigation to create a fully harmonised and trusted system: data hosting, data discovery and checking, data security and processing location. In Table 5, a series of 'nice to haves' is listed. These would be options that are worth considering as they would potentially improve the overall system, but are not currently thought of as blocking the development of use of a data processing system[36].

Table 4: Gaps identified

Topic	Summary
Data hosting	There is a requirement for the data to be hosted, if a collaborative system is to be used and if Google Earth Engine is to be moved away from. Ideally the data would remain with the data providers, but this might not be practically possible. However this issue is solved, reporting countries and agencies need to be able to trust the source of the data.
Data discovery and suitability checking	Discovering the 'correct' data to use in the LDN reporting process is critical. Enabling greater freedom for data input choices requires a robust and managed data discovery process, as well as a suitability checking process against the minimum data quality standards.
Trust and security	If national data are to be hosted on an online platform, or processed on commercial servers then the country analysts need to trust that the data holding, transfer and processing systems are trustworthy.
Options for cloud vs local	Analysts and users should have the option regarding where the national and sub-national data are stored and processed.

Table 5: "Nice to have" options for a LDN processing system

Topic	Summary
Sandbox tools	Being able to assess and develop different datasets and methods is seen as a useful option.
Upfront cost assessment	Understanding the costs of implementing a national or

[36] The topics listed in Tables 4 and 5 have been extracted by the author from the Consultation responses and a synthesis of knowledge gained from work undertaken for WG2 and research into various online resources.

	subnational LDN assessment is important to many countries and being able to provide a range of likely costs based on different implementations would be useful.
Management of processing tools	Being able to create and share models as part of a community ecosystem could help generate new ideas and further a sense of inclusivity in the process.

Data hosting

As part of this review, representatives from GEO and UNCCD were approached and asked for their views on the capability and willingness for those organisations to curate and supply sub-indicator data, either at present or in the future, and whether or not there was the possibility of using servers hosted by these organisations for LDN data processing. The reason for these questions was that GEO and UNCCD are seen as impartial, but interested, parties. It is a possibility that these institutions could have a direct stake in any data processing system, at the same time reducing concerns surrounding impartiality. However, if these organisations are unable or unwilling to take on this role then there is little reason to pursue this.

GEO reported that at present the organisation did not have the capacity to host data on their servers. However, this is a capacity envisaged for the GEO Knowledge Hub which will hopefully be formed after a decision by the GEO Executive Committee in July 2020. They also stated that they have good relations with several cloud service providers, some of which are commercial, but also with public providers such as the Open Science Data Cloud[37]. GEO has the desire to help in the area of data hosting and is looking to develop their capacity.

UNCCD reported that they did not have current capacity for LDN data hosting but are in the planning phase of a new geospatial platform for reporting. The provision of specific data hosting and processing capacity is less of a key consideration for UNCCD whose focus is primarily on developing a reporting platform.

Data sharing

The respondents to the survey mentioned earlier in this report stated that any system development should build on trends.earth without favouring any one specific dataset or algorithm. Submissions into a data processing system should strive to meet the minimum DQS for data providers and include relevant metadata. It has also been stated that a suitable way to share data would be to use an Open Geospatial Consortium (OGC) web service such as the web coverage service (WCS). The Brazil Data Cube[38] (BDC), a subproject of the project “Environmental Monitoring of Brazilian Biomes“, is developing a series of interesting web services around the concept of data cubes and the types of data commonly held within them. These include the Web Time Series Service, the Web Data Cube Service and the Web Land Trajectory Service[39]. All of these should be investigated further for their relevance to an LDN system.

[37] https://www.opensciencedatacloud.org/
[38] http://brazildatacube.org/
[39] https://github.com/brazil-data-cube

On a separate issue, regional hubs could be considered, to provide data to countries in a geographic location with similar requirements, and to minimise the need for repeated processing of the base data by each country. It is acknowledged that political, geographical and technical issues might limit the effectiveness of such an idea.

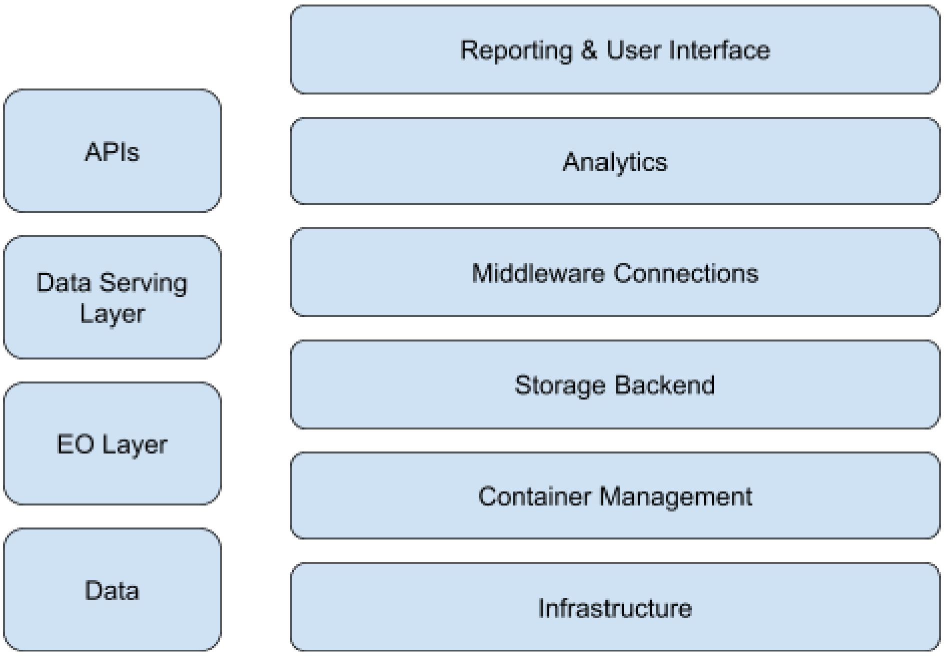
Reporting

Although not falling under the direct remit of this Technical Note, some interesting work is being done on open systems for data reporting for the SDGs. OpenSDG is a platform for managing and publishing data and statistics and is the result of collaboration between the UK Office for National Statistics (ONS), the US government, the nonprofit Center for Open Data Enterprise (CODE) and members of the Open SDG community. The platform helps countries and organisations gather and disseminate national and local data and improves access to official statistics and metadata. This project could be built into any future reporting process, if it is thought to be an advantage over existing methods.

Conclusions

This review has attempted to provide an overview of, and links to, existing issues, technologies, code and data that could be used in the development of a data analysis system. The components of such a system would be as shown in Figure 2. Different tools and data can be used to address different architectural components and some components could be removed if not required.

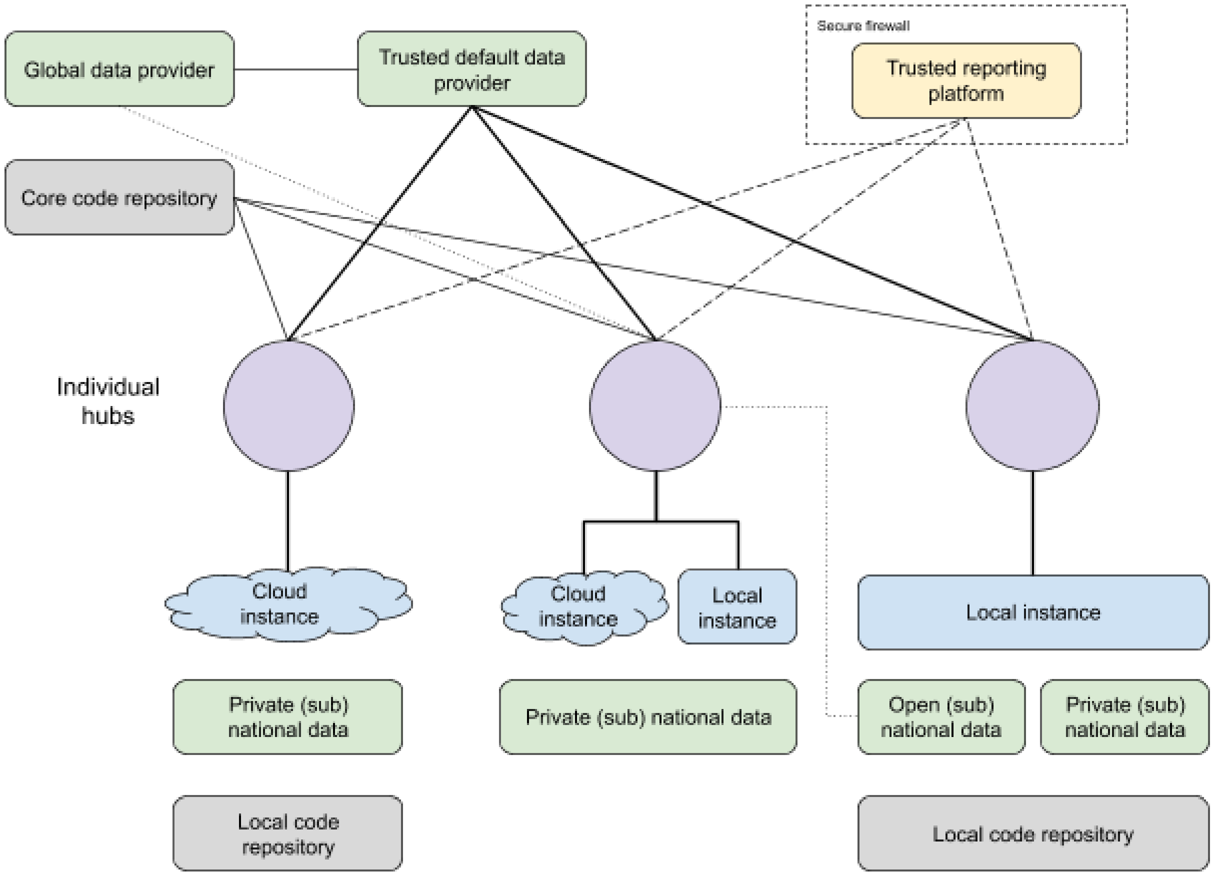
Figure 2: A simplified possible architecture for an LDN data analytics and reporting system[40]



Although this provides a simplified architecture for a data analytics platform in terms of the individual components, it does nothing to explain how the components might interact with each other. One possible implementation of an LDN system could be as shown in Figure 3.

[40] Based on the architecture presented at http://biggis-project.eu/biggis-docs/architecture/arch-overview/

Figure 3: A possible implementation of a federated system



The system in Figure 3 shows three individual hubs (these could be hosted and provided by national governments, organisations or provided as regional hubs involving interested parties from different countries, depending upon requirements). Indeed one of these hubs could even be hosted and managed by the GEO LDN Initiative or some other trusted organisation. Each hub takes in data from either a global data provider or from a trusted data provider/curator. Code is also pulled in from a core repository. Each hub could be run on either cloud services or as a local instance (in some cases a mix of both might be required, and any proposed system should be able to account for this). Importantly, private data and code repositories should sit within the hubs, with the option for sharing to the rest of the network if suitable. For instance, if an efficient method of processing is found, or a meaningful new way of generating data for a sub-indicator, then analysts might want to make this openly available from their repository for use by members of other hubs. Ultimately, all of the hubs then transfer the LDN results in a standardised manner into a trusted reporting platform hosted by the UNCCD.

The creation of a collaborative and modular system is a difficult undertaking, but one that is laudable. Building in an open, stepwise manner, based around the data quality standards from WG 2 and involving WG1 to inform the ultimate users, it is possible that an SDG wide system could be constructed that is flexible and easy to use.

Appendix A: Verbatim Survey Responses

Table A1: What comments do you have regarding the possible creation of a federated analytic platform for SDG 15.3.1 data analysis?	
Gregory Giuliani	I think this can be extremely valuable for the users but there are several important challenges to answer the needs of end-users: they need to find the data that they want to achieve their work; if the system will be distributed, interoperability will be an important issue to solve to ensure that the various systems will adequately talk to each others; we are also talking of probably large volumes of data that might be an issue the move such large volumes and therefore we need to carefully think where the processing will be executed; and an essential point is to have good reference data for training and validation of algorithms/outputs.
Xiaosong Li	A federated analytic platform is important for SDG 15.3.1 analysis, which should provide both SDG report guidance and GIS tools. Of course, the available dataset should be incorporated.
Marc Paganini	Trends.Earth is de facto the official UNCCD tool for monitoring and reporting on SDG 15.3.1 and should be seen as being THE integrator of the federated analytic platform for SDG 15.3.1. This also implies that the SDG 15.3.1 front end (with associated graphical user interface) should be provided by Trends.Earth. To be a federated analytic platform, trends.earth must be inclusive, transparent and participatory. In practical terms this means that trends.earth should not favour a given dataset or a given algorithm. Trends.earth should be a collaborative platform and offer to the users (which means the countries) different solutions in terms of input data and in terms of processing algorithms. This would also be an incentive for data providers and algorithm developers to contribute to the development of a federated analytic platform. This also implies that each new dataset and each new algorithm should be provided in a fully transparent way with all the necessary metadata. These metadata data should help countries deciding which dataset/algorithm to use. Full transparency will also build country confidence and ownership. However trands.earth should not accept any type of datasets or algorithms. There should be some prerequisite conditions for datasets/algorithm on-boarding on trends.earth. Regarding the federated approach, it is important that the data remains with the data curator. We should avoid as much as possibly to make copies of the datasets. Leaving the data by the curator will allow to have the last data collections always available. This implies that the data curators must give on-line access to their data using OGC standard (e.g., WCS, WPS). In this context the use of WCS services would be extremely beneficial since this would allow to get a fast access to the dataset in time and space when needed. In terms of processing, 2 options are possible. Either to do some remote processing on the data provider/curator sites (using WPS standard for example) or to transfer the data sub-set (via WCS standard for example) for processing on the the trends.earth processing environment. The choice will depend on the type of processing (for example a complex processing using different input data will need

	to move data across) and on the agreements with the data provider/curators (on-the-fly data processing services must be offered by the data providers). This means that all the data providers should interface with the Trends.Earth back offices and provide on-line access to their data sets. For what regards the processing algorithms, here also there are 2 options. The data-specific algorithms can possibly remain by the data providers for on-the-fly data processing. This is needed for example for the data pre-processing. Here an agreement must be found with the data providers if these data-specific algorithms are not part of the on-the-fly data processing services offered by the data providers. The second option is to have the algorithms installed on the Trends.Earth environment. Here I would advocate for open-source algorithms as a baseline. The federated system should also offer some interactive tools (Sandbox applications such as Jupyter notebooks) to allow expert users from the countries experimenting and developing their own processing systems. Another important aspects is the availability of auxiliary information (e.g. meteorological data) that would help countries to better access land degradation. Access to auxiliary information should follow the same logic and stay with by the data providers as much as possible. Last point is the possibility for countries to use their national data, which can be sometimes sensitive. This can be done either by using the data locally or by uploading their data on the trend.earth processing environment. Here there is an issue of trust. So trends.earth should provide all the evidence that the data when uploaded is treated with all the privacy and security measures. This also implies that trends.earth must have a local component (desktop application) for the data that the users don't want to upload on the cloud.
Michael Cherlet	We are definitely interested in federated platforms. With the EU Copernicus Global Land we are looking into this as well. We have currently some experts analysing the options for federated open cloud solutions. Through this we could contribute - in the near future, as for now we still operate on diverse computer infrastructure setups.
Brian Killough	A federated system is a good idea as it will allow a diversity of systems and datasets to be applied to the SDG. I would stress that these systems should be free/open to allow transparency into the entire process (data to product) and repeatability of solutions over time. In addition, I would suggest such systems allow global analyses to benefit the largest number of countries.
Emil Cherrington	From my point of view as an Earth observation scientist, the development of a federated analytical platform for SDG Indicator 15.3.1 seems like a laudable goal. I provide some additional feedback in response to the question on hosting, data security, and processing. One would assume that on the processing side of such a platform, that it would be cloud-based. From our experience in SERVIR, working with international centres in different regions, one concern regarding such a platform would be whether or not cost might be a barrier to access. That is, does federated imply that each country or region must develop its own platform, at cost to each country or region? Or could it be such that the platform itself is somewhat centralized (and paid for) and each country or region administers their own part of the platform?

Rob Waterworth	- I think it is vital if we are looking to reduce costs, get more countries involved and stop building lots of simple, cheap systems that do not really meet all our needs. I very strongly support this work. - However, rather than creating a new analytic platform though I would look at current generic solutions such as FLINT(nb conflict of interest given my role with moja global!) and build a collaboration that benefits everyone. - These are not cheap to build and maintain: FLINT already has many millions of $ behind the development, so try not to do this again! - The key is to get collaboration on the build and ongoing management of the platform. Getting this collaboration this takes a lot of time to build from scratch (we have found this out through bitter experience) but once you start to get traction it moves more quickly. As such, try to work with an existing system and build on the experience and collaborative processes. This will reduce development time, increase reliability and use. - Open source will be key, but only if the code is well managed through an existing open source project or organisation. There are lots of different options for doing this: for example moja global is a project under Linux Foundation while the Global Earthquake model is its own organisation. - Please note that open source is not as simple as just making the code available, there is a huge process behind it. It is the management processes that groups like Linux put in place that have made open source so powerful, not just the code itself. - Look to collaborate with other groups working in open source, such as DIAL
Philip Frost	The ODC has provided a unique platform to allow entities to develop customized applications, especially in support of SDG's. The ability to produce a basic Land Degradation product for the region which can be improved over time is more important than waiting for the perfect solution to come. Land Degradation is unfortunately a very contentious topic with a very wide range of opinions. The implementation of the ODC is rather complicated and requires skilled resources as well as significant computer infrastructure. What is currently missing is a dedicated portal to allow organizations to derive Land Degradation using standard algorithms with global inputs.
Mariano Gonzalez-Roglich	To be honest, I keep on hearing about this, but I can not really understand what it means or how it would be implemented. Would this be something similar to SEPAL or data cubes in which there is a common repository and then people can share code in different languages (R, Javascripting, python, etc) to produce SDG related analytics? Specific examples or a more thorough explanation on what a federated analysis platform means would be needed for me to be able to answer this. I fully support open code, sharing information and standardization, so if it goes in that direction, I would say I support it. That being said, someone needs to manage such a platform, maintain it, and coordinate efforts, and that is why details are key.

Table A2: What SDG specific data processing platforms are you developing or aware of?	
Gregory Giuliani	On our side, we have used: the Swiss Data Cube, Google Earth Engine, and the GEOSS platform (with the Virtual Lab.)
Xiaosong Li	We are developing a CASEarth SDG toolbox for SDG reporting, including SDG 2, SDG 6, SDG 11, SDG 13, SDG 14 and SDG 15, based on our own digital earth platform. Also, the code could be applied to Google Earth Engine, data cube, etc.
Marc Paganini	First ESA is developing a Data Cube Servicing Facility called EuroDataCube (https://eurodatacube.com) that provides easy and fast access in space and time to the sentinel data in analysis ready data format (SentinelHub) and to high level data sets or derived data (xcube). It also offer services to bring users data and algorithms on a data cube. The EuroDataCube services follow the OGC standards for full interoperability, which would allow Trends.Earth to use its existing tool set and interconnect with EuroDataCube for data discovery, view and download. Data cubes organise satellite imagery into stacks of consistent, calibrated, geographic 'tiles', underpinned by a relational database ready for rapid manipulation in a high powered, computing environment. Data cubes can potentially serve the EO data and information needs (in terms of data discovery, access, processing and analysis) of the large community of SDG 15.3.1 stakeholders, addressing both global and national data processing needs. As part of the European Copernicus Programme and on behalf of the European Commision, ESA is developing 4 cloud-based platforms known as the Data and Information Acess Services (DIAS) (https://www.copernicus.eu/en/access-data/dias); CREODIAS, SOBLOO, MUNDI and ONDA. An additional platform is developed by EUMETSAT: WekEO. The DIAS systems facilitate access to Copernicus data and services, including the Sentinel data, as well as to on-line EO data processing and analytic tools. The five DIAS online platforms allow users to discover, manipulate, process and download Copernicus data and information. All DIAS platforms provide access to Copernicus Sentinel data, as well as to the information products from Copernicus' six operational services, together with cloud-based tools (open source and/or on a pay-per-use basis). To be noted is the fact that EuroDataCube instances run on the DIAS systems as well as on Amazon Web Services (AWS) ESA is developing Thematic Exploitation Platforms (TEPs) to help users access an interconnected, and virtual work environment, providing access to sectorial data (EO and non EO) and the data processing and analytical infrastructure required to work with them. Instead of the user downloading and working locally on the data, the TEP allows the user to access these sectorial data on the cloud. All TEPs come with their own front end (graphical user interface) but also include man machine interfaces allowing some interoperability with Trends.Earth back office. Currently there are TEPs addressing the following applications: coastal, forestry, hydrology, geohazards, polar, urban and food security. Some of the TEP datasets are very relevant for SDG 15.3.1 such as the World Settlement Footprint (WSF) will provide a global mapping of human settlement. In October 2020, the U-TEP will release the WSF Evolution which will provide annual changes of the human settlements worldwide for the last 30 years.

	ESA will initiate in 2021 the SEN4LDN project to develop and showcase new EO methodologies that exploit the high frequency and spatial resolution of the Copernicus Sentinels observations (Sentinel 1 and Sentinel 2) complemented by Landsat 8, . This will help countries setting their land degradation neutrality targets and reporting on their progress at appropriate scales (management relevant units such as agro-ecological areas or administrative unities). The project will develop some open source software for a robust automated mapping of land cover and its changes and productivity dynamics at high spatial resolution. The SEN4LDN open source software will be available for use on trends.earth.
Michael Cherlet	As we discussed in the GEO-LDN group and with Conservation International (CI), we are developing a stand alone (cloud compatible) tool to implement the 'convergence of evidence' (COE) approach that we launched in the World Atlas of Desertification (WAD) (https://wad.jrc.ec.europa.eu & https://wad.jrc.ec.europa.eu/countryreport). We have had discussions with CI to ensure compatibility with the then planned further developments of the 'trends.earth' capability. The COE allows to combine a numbers of data/information layers in a non-assumptive way as to indicate (in case of WAD) were various issues that relate to land degradation simultaneously are currently active, posing a 'pressure' on the environment. This is mane not directly related to the stipulated SDG reporting, but it allows to either combine the 3 indicators in a controlled but non restrictive way and eventually to add more layers to this for further/deeper/more local interpretation. This tool is now newly ready and I think we can have another meeting with CI after summer.
Brian Killough	As you know, we founded and use the Open Data Cube (ODC). We have demonstrated operation on the Amazon Cloud connected to AWS datasets and on the Google Cloud connected to Earth Engine datasets. We also have a number of open source algorithms and utility scripts in Jupyter notebooks.
Emil Cherrington	In the context of SERVIR, and specific to SDG Indicator 15.3.1, our colleagues at the Nairobi-based Regional Centre for Mapping of Resources for Development (RCMRD) have been doing modelling related to land productivity and land degradation, in support of RCMRD's Member States in Eastern & Southern Africa. Among other things, one of the platforms they have been using for data processing – and data sharing – to make the modelling process interactive is Google Earth Engine (GEE). (That is to say, RCMRD has been using GEE to develop its platform / system for tracking land degradation.) Besides outright mapping of degraded lands, RCMRD has also been modelling the impact of invasive species in their geographic domain, as these also have an impact on land productivity. They recently published a study (Ouko et al. 2020) on the impacts of invasive species in Kenya's rangelands, noting the relevance to SDG 15. Relevant to the sub-indicator on land cover, RCMRD also has a land cover monitoring service, and has done work mapping forest degradation in the context of their Kenya Water Towers monitoring work. Anastasia Wahome (awahome@rcmrd.org), the SERVIR Eastern & Southern Africa Science & Data Lead at RCMRD, can provide additional details on RCMRD's LDN-related activities. Also in the context of SERVIR, Dr. Amadou Dieye of SERVIR-West Africa consortium member Centre de Suivi Ecologique (CSE), who has participated in GEO's LDN Working Group on Data Analytics, has called for developing a West

	Africa-wide system for tracking relevant SDGs using Earth observations, but we are not aware of the current status of the proposed activity. Also perhaps relevant to the processing platform issue is to mention that in 2005, in Belize, we used ArcGIS to do simple modelling of (overlay analysis) of areas likely degraded, in combination with interviews done with key informants (Meerman & Cherrington, 2005). In 2019, at their request, we provided an update on potentially degraded agricultural lands across Belize to the Forest Department / Ministry of Forestry, Fisheries, the Environment, and Sustainable Development (MFFESD). That update was also done using ArcGIS.
Rob Waterworth	● There are numerous GIS type platforms (MapX, etc etc): but many are quite limited in general ● However, there are very few model-based data analysis platforms. The FLINT being the only one I can readily identify (this gap is whey we started building it)
Philip Frost	We are in the process of setting up a data cube for Syria, Jordan and Lebanon based on the system provided by CEOS and supported by the Swiss and Australian cubes. The Jupyter notebook for the production of SDG 15.3.1 was implemented as published by CEOS on the ODC portal which is a simplistic version of the trends.earth product. The Landsat 8 archive available on the AWS cloud from 2013 - 2020 served as input for the calculation of SDG 15.3.1. Once the Landsat level 2 global data set becomes available we will be able to apply the Land Degradation product further back in time which will allow for more accurate results. One of the biggest limitations within the current CEOS Land Degradation application is the use of the ESA 300m global Land Cover product. This product is not sufficient for accurate land degradation assessments across the Middle East due to its low resolution. We are in the process of integrating the Copernicus 100m global Land Cover from VITO which should provide improved detection accuracy for smaller land cover classes. The CEOS implementation also does not include any climate data which is essential in order to assess true Land Degradation. [41]iMMAP also provided information on the creation and use of an Open Data Cube for use in assessing urban environments (SDG 11) in the Middle East. Multiple datasets (MODIS, Landsat, Sentinel 2, VIIRS and very high spatial resolution commercial imagery) were used in this work, as were a number of spectral index algorithms. The data cube was deployed on AWS and also calculated SDG 15.3.1 and 6.6.1 indicators.
Mariano Gonzalez-Roglich	Conservation International leads the development of Trends.Earth, a tool for computing indicators for SDG targets 15.3.1 and 11.3.1.

[41] This text is provided by the report author, based on documents supplied in place of direct responses to the survey questions.

Table A3: Are you aware of any federated analytic platforms related to the other SDGs?	
Gregory Giuliani	No
Xiaosong Li	CASEarth SDG toolbox mentioned above is a specific platform for different SDGs. Also, EO4SDG, UNEP and GEO have contributed some tools, such as SDG 11.3.1, SDG 15.3.1, SDG 6.6.1.
Marc Paganini	The GEO EO4SDG initiative has launched the concept of the GEO SDG toolkits. See concept note in attachment. But these toolkits are meant to be knowledge hubs on SDGs rather than analytic platforms. GEO is developing a GEO Knowledge Hub for Open Science Solutions to Societal Challenges. The GEO Knowledge Hub will allow some data processing and analytics. The ESA TEPs are evolving into federated sectorial platforms providing data processing and analytics capabilities and remote access for specific data access and processing. For example the Urban TEP will offer some integrated applications for SDG 15.3.1 which will combine different datasets and algorithms, including the use of a Jupyter notebook for experimental development.
Michael Cherlet	yes, the JRC produced the Global Surface water Explorer (based on 38 years of Landsat data) jointly with Google Earth engine. There in a tool in GEE to explore this dataset (https://global-surface-water.appspot.com). Under the Copernicus Global Land we produce a Lake and River Water Quality product (https://land.copernicus.eu/global/products/lwq) and JRC set up a specific platform for UNEP to serve the SDG 6.3.2. (https://communities.unep.org/display/sdg632).
Brian Killough	Some examples are the ESA Thematic Exploration Platform (TEP) and the JRC Water Product (SDG 6.6.1).
Emil Cherrington	I am not personally aware, but Dr. Argie Kavvada of NASA (argyro.kavvada@nasa.gov), who serves as the Executive Secretary of GEO's Earth Observations for the Sustainable Development Goals (EO4SDG) initiative likely has a better purview on such analytical platforms.
Rob Waterworth	There are a few around climate change: SEPAL from the FAO is great for cover change work for example.
Mariano Gonzalez-Roglich	No

Table A4: What issues around hosting, data security and processing do you foresee with such systems?

Gregory Giuliani	Large volume of data; data sovereignty > would a country agree to share some of their data on cloud-based platforms? For example in Switzerland the government do not wish to have some national data stored on US-based cloud platforms (for obvious reasons of security); in terms of processing now that users are familiar with the high capabilities of platforms like Google Earth Engine or Amazon, they might expect such a level of performance (and that will be challenging). In terms of hosting, maybe an institution that offers neutrality (like the UN-system) might be a good idea.
Xiaosong Li	Data consistency and processing capacity are two key problems.
Marc Paganini	Some of my comments above could be moved to this section. Data hosting implies that trends.earth must find an agreement with the data providers/curators to get remote access to the datasets and possibly do specific on-the-fly processing on the data providers host. In practical terms it means that trends.earth would delegate the data curation and access to the data providers. To give a practical example, the access to the World Settlement Footprint could be provided to Trends.Earth by remotely accessing the Urban TEP through OGC standards (e.g. WCS). Regarding the country owned data, I would strongly encourage a mix between national data used locally (on the desktop part of trends.earth) and data uploaded (on the cloud processing part of trends.earth). For example all the national data that are utilised only for visual analysis to help the land degradation assessment should stay locally. Only the data needed for massive processing should be uploaded.
Michael Cherlet	I am not expert in this really, but one of the 'problems' is the cost of course. Not so much of the hosting but performing analytics on a (federated) cloud system has the cost of processing, but also a considerable egress cost. This is e.g. one of the aspects we want the experts to investigate for our Copernicus service. I don't think security is an issue as such a platform would put publicly available data at disposal and countries would use their own data 'at home'.
Brian Killough	Users should be given choices for hosting and data management so they can determine the best solution for their situation. Cloud computing seems to be the cheapest method. As for data security, this is always an issue. The primary requirement seems to be the need to include local (non-open) datasets (e.g. land class maps) and storage of products in a secure location. Processing seems to be improving every day with cloud computing, so parallel processing and machine learning will be needed for the future.
Emil Cherrington	I don't know that I can provide any insights into the data security of systems, but I can share some insights relevant to hosting, and processing from some internal discussions we have had with representatives from across our SERVIR network of five hubs in Africa, the Americas, and Asia. In terms of hosting, if the idea is to have a federated (vs. centralized) system, then that is congruent with feedback that SERVIR hubs have often received from their member governments, where there is a need to strengthen national capacities and allow for countries to manage their own platforms which could be interconnected in a federated system. That is probably also relevant to data security. In terms of processing, relevant to the discussion is that SERVIR hubs have increasingly turned to cloud computing for data processing, and depending on the volumes of input data for land degradation modelling, in some regions with limited Internet bandwidth, it

	might indeed be necessary for such data processing to occur on the cloud. While not specific to SDG Indicator 15.3.1, the overall trend seen by SERVIR hubs in their respective regions is that geospatial processing in general is moving to the cloud. For instance, SERVIR-West Africa participated in and co-sponsored the African Geospatial Data and Internet Conference in October 2019, and the issue of the necessity of cloud computing was front and centre. That said, however, another consideration regarding accessibility of cloud-based processing is the accessibility of such resources (i.e. cost), given budgetary constraints for many governments. One hopes that the cost of cloud computing does not become a barrier to implementation regarding work on SDG 15.3.1.
Rob Waterworth	● Ideally the platform itself is open source, but allows for lots of different methods of hosting and use. For example, with FLINT there are three main options: go to a commercial platform (FLINTpro), pay a groups to build you r own system (FLINT-based) or hire your own team and use the open source systems (FLINT). This allows a wide range of organisations to use the same core system based on their needs, capabilities and funding. ● Security etc then depends on the option the group chooses to use… ● Processing will be an issue, and the cloud has obvious advantages over other systems ● Also note the issue of open source licences. There are a large number and it is important to pick the right one that fits the needs of multiple groups ○ Moja chose MPL2 so that governments and organisations could use and contribute to the FLINT without having to also contribute all their models and other systems. This means we can offer lots of different use options (govt, self run, commercial). Other licences are much stricter on this limiting options. ○ Choosing the wrong licence can lead to real problems down the track. ● Also note limitations in current UN processes and contracts for open source projects ○ Happy to discuss this, but it is a real problem.
Mariano Gonzalez-Roglich	Again, without a more clear understanding on what it means and the structure it would take, it is hard to comment. Global EO datasets and code can be easily and freely accessed and shared, so I would not see initial problems there. Data property and security becomes an issue if this platform intends to also host/manage datasets produced and owned by countries.

Table A5: What SDG 15.3.1 specific analytical tools are you developing or aware of, and how will they add to the monitoring and reporting process? These might be GIS tools, Jupyter notebooks, dedicated software, Earth Engine javascript etc.	
Gregory Giuliani	We are using Trends.Earth as a GIS tool; we also ported some code on the Swiss Data Cube (mostly for Land Productivity and use Landsat/Sentinel-2 data) and code is available as a Jupyter Notebook and we plan to develop

	some code on GEE for the European Environment Agency.
Xiaosong Li	We are developing a SDG 15.3.1 reporting tool with Earth Engine javascript, which has the capacity of SDG 15.3.1 monitoring and reporting at pixel and regional scale. It is worth noting, providing the global consistent SDG 15.3.1 reporting map in different scenarios, such as climate correction, different land cover etc. could help countries self-reporting and global methodology refinement, which is better than only providing a guidance on how to select dataset and analytical methods.
Marc Paganini	Jupyter notebooks are needed for experimental development and analysis by expert users. Some graphical interactive tools should also be provided for data visualisation, statistics and analytics. A good example is the Visualization and Analysis Tool (VISAT) of U-TEP that allows experts and non- experts to conduct geospatial data visualisation and analyses including on the fly statistical analyses. VISAT is also available as open source and could potentially be integrated on trends.earth.
Michael Cherlet	See above on the COE tool (we might make this also available as QGIS plugin and maybe as notebook). In the near future the Copernicus Land service will also provide more notebook solutions for e.g. EO product time series analysis etc.
Brian Killough	We have been working on a specific Jupyter notebook for 15.3.1. This notebook does not provide the full solution including carbon data, but it does address land change and land classification comparisons to known datasets. When comparing our work to other 15.3.1 tools, I do not feel we come close to the complexity and rigor provided by other groups, but our work could be a catalyst for more development.
Emil Cherrington	My knowledge related to SDG Indicator 15.3.1 may not be so current. Nevertheless, where updated land cover data are a key input for land degradation monitoring, there are many resources available, although those resources likely have to be curated. I would also assume that the GEO LDN Working Group on Data Quality Standards might be providing guidance on standardizing methodologies related to this. In terms of resources, I am aware of many Google Earth Engine and associated Google Colab notebooks developed for updating of land cover datasets, and which are also relevant to the sister GEO initiative, GFOI.
Rob Waterworth	As noted above, FLINT is not specific to this, but being modular means it could do a lot of the work here easily BUT: it does not replace good data collection etc – all this needs to be done still
Mariano Gonzalez-Roglich	Trends.Earth is the tool we developed and maintain. It supports computing SDG 15.3.1 sub-inidicators (changes in land cover, soc and productivity) using global datasets and also allows users to perform the same analysis with their own datasets if they prefer (usually higher spatial resolution). Website: http://trends.earth/ Code repository: https://github.com/ConservationInternational/trends.earth

Table A6: What new datasets and/or analytical code are you aware of specifically related to the three sub-indicators?	
Gregory Giuliani	Except for the code we are developing and some work we are doing with a colleague in Bolivia (self-adjusting algorithm for Land Productivity), we are not aware of other code or datasets. We are planning to develop some code using AI and Landsat/Sentinel-2 data for Land Cover - Land Cover Change mapping for Switzerland to get an annual Land Cover map.
Xiaosong Li	CASEarth and its partner have global 30m land cover datasets from 2000-2018, which would be a valuable datasource for SDG 15.3.1 monitoring. Also, we are developing the land productivity assessment code in order to reduce the effects of climate factors.
Marc Paganini	See Sen4LDN
Michael Cherlet	The Land Productivity Dynamics data set that we produced for the WAD and delivered to UNCCD for the first reporting phase, has been at the basis and been driving the selection of the SDG 15.3.1 indicators. the time series went until 2013 and due to lack of resources we didn't fully manage to update the product. However, now we the processing chain ready (inclusive the phenological analysis of vegetation indices times series and the 'post' processing into a land productivity dynamics product (this product consider the dynamics over a certain period rather that an absolute measure or change of the productivity itself. I hope that over the summer now we can process this and update the product until 2019 (still some issues as the phenology is extremely computing demanding and needs really big data systems to process). Also the Copernicus Global Land produces now an operational and quality ensured Global Land cover mapping at 100m resolution - this includes also change maps. the 2015 maps are available and the 2016, 17, 18, 19 maps will be published during August this year. This constitutes a higher resolution option that the ESA CCI maps (which are mainly made to satisfy the climate community) and continuity is ensured (https://lcviewer.vito.be)
Brian Killough	Most SDGs do not use radar data for their analyses. These data are free from cloud contamination and likely provide a good resource for improved land change detection and even land classification. The key to using these data is preprocessing. Our NASA-CEOS team is developing a Sentinel-1 data cube generation tool to prepare analysis-ready data into data cubes. Such data may have high value to many users.
Emil Cherrington	As indicated, there is a lot of analytical code (and data) on the land cover sub-indicator, which would be relevant, but I am not aware of a single curated source of such data. Probably such an undertaking would be worthwhile.
Rob Waterworth	In the remote sensing world this is rapidly and ever evolving: the challenge is to find the good robust ones… I am sure you are more on top of this than me!
Mariano Gonzalez-Roglich	As part of the Tools4LDN project (https://www.tools4ldn.org/) we'll be adding functionalities to compute productivity with 30 m landsat/sentinel harmonized data for monitoring progress towards LDN. If global land cover datasets become available, those will also be added into Trends.Earth

Table A7: What analytical code are you aware of that might be being developed for other SDGs that could be forked and altered to be made suitable for SDG 15.3.1 monitoring and reporting?	
Gregory Giuliani	None
Xiaosong Li	To my knowledge, GEO SDG 11.3.1, UNEP SDG 6.6.1 both could be forked and altered for SDG 15.3.1.
Marc Paganini	ESA will initiate in 2021 the SEN4LDN project to develop and showcase new EO methodologies that exploit the high frequency and spatial resolution of the Copernicus Sentinels observations (Sentinel 1 and Sentinel 2) complemented by Landsat 8, . This will help countries setting their land degradation neutrality targets and reporting on their progress at appropriate scales (management relevant units such as agro-ecological areas or administrative units). The project will develop some open source software for a robust automated mapping of land cover and its changes and productivity dynamics at high spatial resolution. The SEN4LDN open source software will be available for use on trends.earth. Objectives are: ● Develop a robust automated EO method to map land cover and its changes at a high spatial resolution for large-scale applications (robust change detection for LDN) ● Develop a robust automated EO method to monitor land productivity dynamics at high spatial resolution (robust phenological analysis of vegetation parameters for LDN) ● Assess the feasibility of deriving information on carbon stocks above and below ground, or any related proxy (e.g. SOC) ● Develop an integrated EO-based solution to assess land conditions (and in particular land potential and resilience) at high spatial resolution ● Examine methods to integrate the 10-30m datasets with the existing default 300m-1km data streams ● Demonstrate adequacy of approach by deploying the proposed method in a number of biomes that are representative of the diversity of land degradation processes. ● Organise an international round-robin comparison of EO methods ● Implement the method on an open-source and free-of-charge software tool with all necessary processing workflows, and integration on cloud-computing platforms. In particular the project will analyse the usefulness of the Sentinel 2 / Landsat 8 fused products (Level-2F) at 10m spatial resolution.
Brian Killough	I am not aware of any other activities.
Emil Cherrington	Similar response to the previous question.
Rob Waterworth	Good question: nothing specific. But if it is possible not to fork and simply add to that would be of greater benefit…

Mariano Gonzalez-Roglich	I am not particularly aware of any. I would imagine that SDG 15.1 would be one in which some tools may be available.

<table>
<tr><th colspan="2">Table A8: What other comments do you have?</th></tr>
<tr><td>Rob Waterworth</td><td>

I think this is a great opportunity to start to look at collaborating on the core analysis frameworks.
The key will be to separate the following things as they often get confused and confounded. This is very problematic and so trying to get clarity around these is important. For example, people often look at outputs and don’t like the results, but then can’;t easily find if this is because of the input data, the models or the overall framework. Ideally the framework is super flexible, the models are configurable, the data can be modified and updated easily, the UIs can be changed as needed for different users.

Framework: the generic system that pulls everything together. Ideally can be used by the entire community
Models: the algorithms etc that drive the calculations: the same models may be used by multiple groups, but not everyone as they can be situation specific
Data: the inputs and outputs: these will likely vary considerably by user due to different conditions and needs
Interfaces: what we all see: there will be generic components but should be easily skinned and changeable.

Ideally any system should also have APIs that allow it to be used by other existing applications, linked to existing reporting system etc

Happy to discuss this more as needed.

Building collaborations between existing tools will also lead to better and faster results, but will require work

We have tried to do this in many cases, with varying degrees of success. To be honest, it is the UN agencies that are the hardest to deal with here…

</td></tr>
</table>