



[www.colandrapp.com](http://www.colandrapp.com)

*Training event*

# Who are we?



Samantha Cheng, PhD  
Assc. Director for Conservation Evidence @  
ASU-CBO



Caitlin Augustin, PhD  
Director of Product for DataKind

# What are we covering today

- The genesis of colandr and where it fits in the space of open source tools
- The process for joining colandr and creating a review
- The colandr community
- Next steps in the colandr user and research community

Any issues, email  
[colandrteam@gmail.com](mailto:colandrteam@gmail.com)

# Goals of the session

- Gain familiarity with the colandr tool
- Feel comfortable creating a review in colandr
- Ask questions

Getting started

# Getting started



[www.colandrapp.com](http://www.colandrapp.com)

How-to docs, FAQs, updates, message board - [www.colandrcommunity.com](http://www.colandrcommunity.com)

Introductions

Make sure you sign-in

Download training materials  
(optional)

# EVIDENCE GAP MAP

Need a better  
way to get  
through piles of  
citations than  
manual sorting

35,000 potentially relevant hits  
from peer-reviewed articles & grey literature



*Screening titles and abstracts*



~3,000 relevant articles



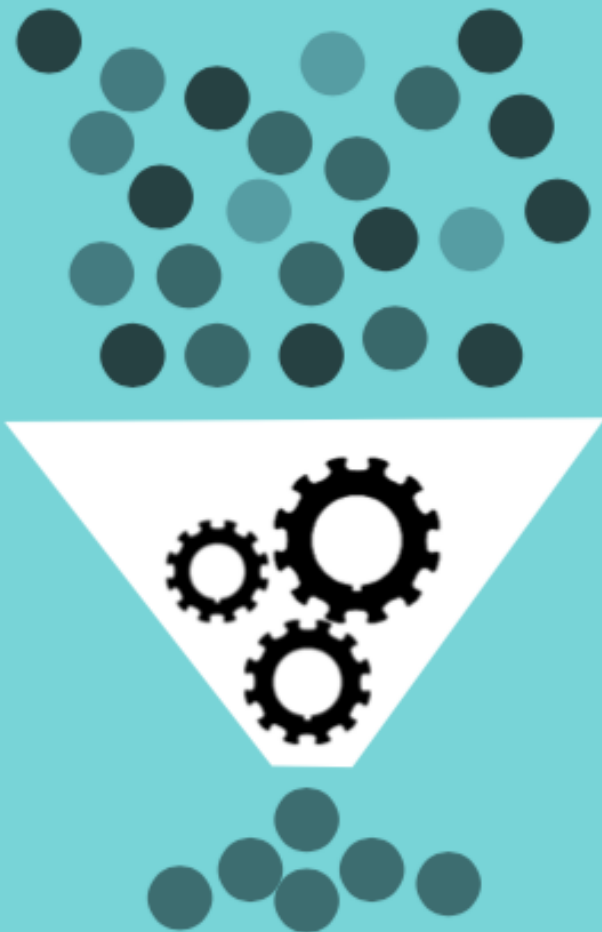
*Reading full-text, extracting meta-data*



1,038 included articles

Machine learning can help  
make evidence synthesis  
**faster** and more **efficient**

Evidence-Based Conservation Working Group








Letter | [Open Access](#)

## Using machine learning to advance synthesis and use of conservation and environmental evidence

S.H. Cheng , C. Augustin, A. Bethel, D. Gill, S. Anzaroot, J. Brun, B. DeWilde, R.C. Minnich, R. Garside, Y.J. Masuda, D.C. Miller, D. Wilkie, S. Wongbusarakum, M.C. McKinnon

First published: 12 April 2018 | <https://doi.org/10.1111/cobi.13117>

**Article impact statement:** Machine learning optimizes process of systematic evidence synthesis ... [More](#) ✓  
This article has been accepted for publication and undergone full peer review but has not been ... [More](#) ✓



A machine learning powered app designed to facilitate evidence synthesis

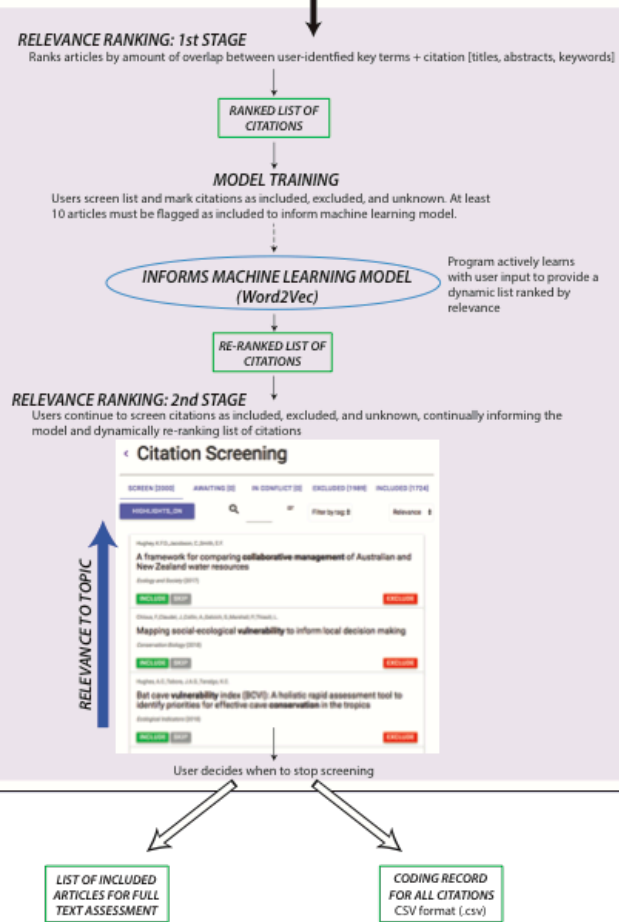
	SYSTEM 1	SYSTEM 2
SYNTHESIS STAGE	Screening @ title/abstract	Data extraction
APPROACH	Relevance ranking of citations	Semi-automated labeling of articles

output

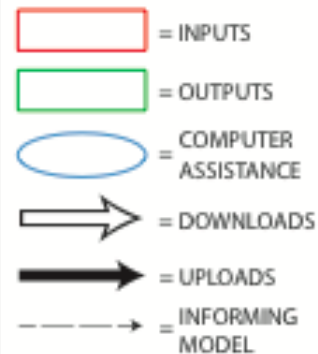
colandr

input

## SYSTEM 1: CITATION SCREENING



## LEGEND



## SYSTEM 2: DATA EXTRACTION

ARTICLES FOR FULL TEXT ASSESSMENT  
Portable document format (.pdf)

### MODEL TRAINING

User manually categorizes at least 25 articles to train model



FOR EXAMPLE:

WHAT CONSERVATION INTERVENTION(S)  
IS THIS ARTICLE ABOUT?

Land/water protection  
Land/water management  
Species management  
Resource management  
Education/awareness  
Law & policy  
Economic incentives  
External capacity building

WHAT HUMAN WELL-BEING OUTCOME(S)?

Economic living standards  
Material living standards  
Health  
Education  
Social relations  
Security & safety  
Governance  
Subjective well-being  
Cultural/spiritual values  
Freedom of choice/action

INFORMS MACHINE LEARNING MODEL  
(GloVe)

DISPLAYS SENTENCES RELATED TO PREDICTED CATEGORIES

### Data Extraction - Label Review

Local impacts and responses to regional forest conservation and rehabilitation programs in China's northwest Yunnan province

Sentences related to Intervention, law & policy

#### 1. Confidence: High

In 2000, the government declared a ban on logging in natural forests (Zhou et al., 2002), as part of the NRP. The NRP's inception actually dates back to the 1970s and 1980s, when the Chinese government developed shelterbelt afforestation programs along major watersheds. In the 1980s, a striping ban system was introduced in response to huge increases in natural wood consumption. These measures increased forest cover to...

#### 2. Confidence: High

They are assigned to the office by the government, and forestry agencies have no choice in it. Only a third of the forestry officials have a prior background in forestry. On the basis of the community survey and discussions with the agency staff, several recommendations were made. These included recommendations regarding: (1) program management; (2) establishing feedback mechanisms across different administrative levels.

#### 3. Confidence: High

Funds are also provided to the forestry agencies to cover technical support and program design. At the case studies described below reveal, these funds are often insufficient (see also Zou, 2012). Table 1 Local government agencies function roles in the GLDP Agency Responsibility Planning Commission (local planning and compiling Ministry Nationality Affairs Commission Mediating in disputes with potential ethnic implications Department of...

User accepts or rejects predicted categories

### Data Extraction - Label Review

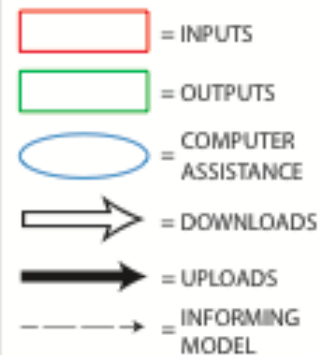
Local impacts and responses to regional forest conservation and rehabilitation programs in China's northwest Yunnan province

Conservation governance in Nepal: protecting forest biodiversity and people's livelihoods	ACCEPTED
Conservation governance in Nepal: protecting forest biodiversity and people's livelihoods	ACCEPTED
Conservation governance in Nepal: protecting forest biodiversity and people's livelihoods	ACCEPTED
Conservation governance in Nepal: protecting forest biodiversity and people's livelihoods	ACCEPTED
Conservation governance in Nepal: protecting forest biodiversity and people's livelihoods	ACCEPTED
Conservation governance in Nepal: protecting forest biodiversity and people's livelihoods	ACCEPTED

	A	B	C	D
1	article id	Author(s)	Conservation_type	Human Well-being type
2	1	Smith et al.	protected areas	economic well-being
3	2	Wang et al.	PES	health
4	3	Gomez et al.	CLIMED	rights

DESIRED DATA FROM  
INCLUDED ARTICLES  
Comma-separated values (.csv)  
Excel & R compatible

### LEGEND



	Software name	Setting up the review	Scoping/pilot study	Literature searching	Duplicate checking	Article screening	Data coding	Critical appraisal	Synthesis	Documentation
	E.g.	Facilitation of question formulation and/or stakeholder engagement	Protocol development , PICO* elements specified	Software integrated with publication databases	Automated marking of duplicates	For study selection	Tagging and extraction to support meta-analyses	Risk of bias assessments	Facilitates quantitative/ qualitative syntheses of results	Output of text, figures or tables to assist with report writing
	CADIMA									
*	Colandr									
	Covidence									
	DistillerSR									
	EROS									
*	EPPI-Reviewer 4									
	HAWC									
	METAGEAR package for R									
	PARSIFAL									
*	Rayyan									
	REviewER									
	RevMan 5									
	RevMan Web	Data unavailable								
	SESRA									
*	SLR-Tool									
	SLuRp									
	SRDB.PRO									
	SRDR									
	StArt									
	SUMARI									
*	SWIFT-Review									
	SyRF									
	<b>TOTAL</b>	5	10	13	11	20	19	12	15	13

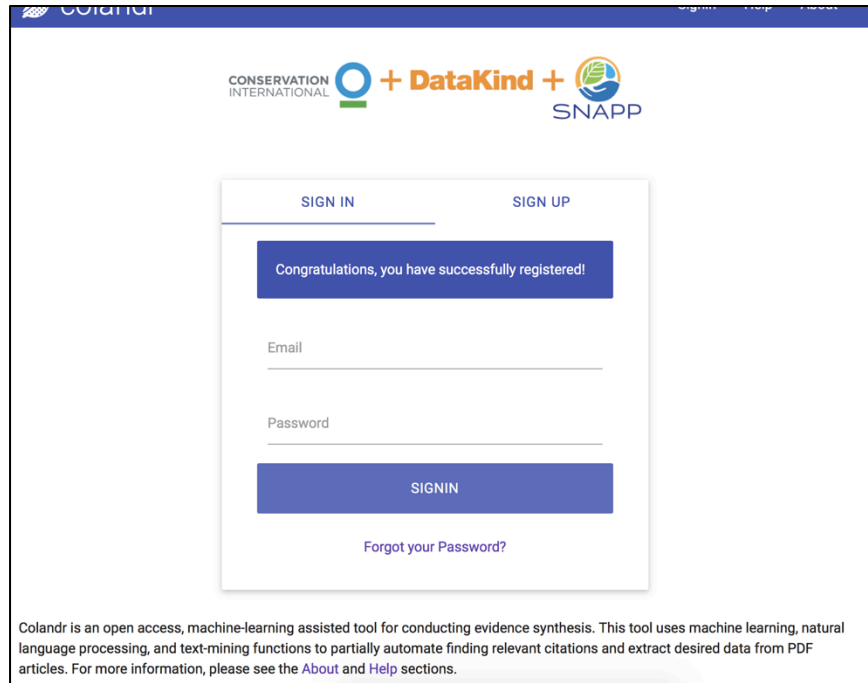
Using colandr

# Part 0: Signing up for colandrapp.com

Did you sign up? If so, great!  
If not ... in this section, you'll sign up for the tool

# Key features

- Will get a confirmation email and must confirm within 24 hours
- Can reset password
- Must have an account to be a collaborator



The screenshot shows the Colandr web application interface. At the top, the navigation bar includes the Colandr logo and links for Sign In, Help, and About. Below the navigation bar, the logos for Conservation International, DataKind, and SNAPP are displayed. The main content area features a central white box with a blue border. Inside this box, the 'SIGN IN' tab is selected, and a blue message box states 'Congratulations, you have successfully registered!'. Below this message are input fields for 'Email' and 'Password'. A blue 'SIGNIN' button is positioned below the password field, and a link for 'Forgot your Password?' is located at the bottom of the box.

Colandr

CONSERVATION INTERNATIONAL + DataKind + SNAPP

SIGN IN SIGN UP

Congratulations, you have successfully registered!

Email

Password

SIGNIN

[Forgot your Password?](#)

Colandr is an open access, machine-learning assisted tool for conducting evidence synthesis. This tool uses machine learning, natural language processing, and text-mining functions to partially automate finding relevant citations and extract desired data from PDF articles. For more information, please see the [About](#) and [Help](#) sections.



# Part 1: Establishing a review

In this section, you'll hit "create" on a review and input some information to get a feel for the planning phase

# Key features

- Guided through review set up
- Have change to develop Boolean search string
- Establish selection criteria and data extraction rules ... without these, you won't be able to progress successfully!

## Boolean Search Query

("conservation" OR "conserve" OR "conservancy" OR "protect\*" OR "management" OR "awareness" OR "law\*" OR "policy\*" OR "reserve\*" OR "govern\*" OR "capacity-build\*" OR "train\*" OR "regulation" OR ("payment for ecosystem services" OR "PES") OR ("ecotourism" OR "eco-tourism") OR "sustainable use") AND ("marine" OR "freshwater" OR "coastal" OR "forest\*" OR "ecosystem\*" OR "species" OR "habitat\*" OR "biodiversity" OR "sustainab\*" OR "ecolog\*" OR "integrated" OR "landscape" OR "seascape" OR "coral reef\*" OR "natural resource\*") AND (("wellbeing" OR "well-being" OR "well being") OR "ecosystem service\*" OR "nutrition" OR "skill\*" OR "empower\*" OR "clean water" OR "livelihood\*" OR "food security" OR "resilience\*" OR "vulnerability" OR "social capital" OR "attitude\*" OR "perception\*" OR "human health\*" OR "human capital" OR ("traditional knowledge" OR "TEK")) AND ("human\*" OR "people" OR "person" OR "community\*" OR "household\*" OR "fisher\*" OR "collaborative")

# Two views - review owner and review collaborator

- The review owner is the only one who can establish and edit the planning phase of a review
- Only the owner can delete a review and edit the settings of a review
- *WARNING: If you change your review from 2 to 1 reviewers required per citation/ article, ensure that there are not articles “left hanging”*

### Basic Information

Review Name  
Conservation-HWB

Description  
McKinnon et al. 2016

✓ UPDATE CANCEL

### Collaboration

Sam Cheng	cheng@nceas.ucsb.edu	Is Owner
caitlin augustin	caitlin@datakind.org	<a href="#">MAKE OWNER</a> <a href="#">REMOVE</a>

User's Email

➤ INVITE

**Screeners per Citation** ☒ One ☐ Two

**Screeners per Full-Text** ☒ One ☐ Two

✓ UPDATE COLLABORATION SETTINGS

### Freeze Review

You may freeze this review at any time to prevent further additions or changes from being made. This change is reversible.

[FREEZE](#)

## Part 2: System 1 screening

In this section, you'll upload a set of test citations

# Key features

- Data upload: .RIS, .BIB, . TXT
  - Imported citations cannot be deleted
  - Max single import size is 40 MB (~15-20k citations)
- Review tagging
- Resolution needed between reviewers
- Learns after 10 reviews
- Colandr will not automatically include/exclude articles or extract data - the user must make final decisions

FROM FILE WITH A FORM IMPORT HISTORY

Import from UPLOAD

Data Source -- Select One --

Status not screened

Source Name Source URL

IMPORT

Hughey, K.F.D., Jacobson, C., Smith, E.F.

A framework for comparing **collaborative management** of Australian and New Zealand water resources

Ecology and Society (2017)

INCLUDE SKIP

Chloups, F., Claudet, J., Collin, A., Gelcich, S., Marshall, P., Thiault, L.

Mapping social-ecological **vulnerability** to inform local decision making

Conservation Biology (2018)

INCLUDE SKIP

EXCLUDE

- ☐ Useful
- ☐ Excluded intervention
- ☐ Excluded outcome
- ☐ No impact measured
- ☐ No comparator
- ☐ Not human population

cancel ok

<https://tinyurl.com/samplecolandrfiles>

# New synthesis or updating one

- Colandr can learn relevance from “scratch” where it learns as you start including/excluding citations
- Colandr can also learn from a training set of data - e.g. a set of included articles from a previous systematic evidence synthesis and use to rank a new search by relevance.

The screenshot shows the 'FROM FILE' tab of the Colandr interface. It features three tabs: 'FROM FILE' (selected), 'WITH A FORM', and 'IMPORT HISTORY'. Below the tabs, there are three input fields: 'Import from' with an 'UPLOAD' button, 'Data Source' with a dropdown menu showing '~ Select One ~', and 'Status' with a dropdown menu showing 'not screened'. To the right of the 'Data Source' field, there are labels for 'Source Name' and 'Source URL'. An 'IMPORT' button is located at the bottom right of the form.

The screenshot shows the Colandr article selection interface. It displays two articles with their titles, authors, and journals. Each article has 'INCLUDE' and 'SKIP' buttons. A third article is partially visible at the bottom. On the right side, there is a legend with checkboxes for 'Useful', 'Excluded intervention', 'Excluded outcome', 'No impact measured', 'No comparator', and 'Not human population'. Below the legend are the labels 'cancel' and 'ok'.

Hughey, K.F.D., Jacobson, C., Smith, E.F.  
A framework for comparing **collaborative management** of Australian and New Zealand water resources  
*Ecology and Society* (2017)  
INCLUDE SKIP

Chious, F., Claudet, J., Collin, A., Gelcich, S., Marshall, P., Thiault, L.  
Mapping social-ecological **vulnerability** to inform local decision making  
*Conservation Biology* (2018)  
INCLUDE SKIP

☐ Useful  
☐ Excluded intervention  
☐ Excluded outcome  
☐ No impact measured  
☐ No comparator  
☐ Not human population  
cancel ok

## Part 2: System 2 reviews+data extraction

# Key features - full text review

- Full text can only be evaluated for papers that have passed System 1 screening
- Colandr will not automatically include/exclude articles or extract data - the user must make final decisions
- Imported papers cannot be deleted either
- Batch download of pdfs script (R) by Stephen Wood on Github
- Papers must be imported one-by-one

Chen, Min, Yang, Naiyi, Zhang, Endi

**Attitudes towards wild animal conservation: a comparative study of the Yi and Mosuo in China**

*International Journal of Biodiversity Science, Ecosystem Services (&) Management* (2010)

[VIEW FULL TEXT](#) [INCLUDE](#) [SKIP](#) [EXCLUDE](#)

Braga, H.J., Brites, R.M., Caramori, P.D., Garrastazu, M.C., Higa, R.C.V., Radin, B., Sousa, V.A., Wrege, M.S.

**Climate change and conservation of Araucaria angustifolia in Brazil**

*Unasylva* (2009)

[ADD FULL TEXT](#)

< Attitudes towards wild animal conservation: a comparative study of the Yi and Mosuo in China

[INCLUDE](#) [EXCLUDE](#)

**Attitudes towards wild animal conservation: a comparative study of the Yi and Mosuo in China**

Naiyi Yang, Endi Zhang\* and Min Chen

*School of Life Science, East China Normal University, Shanghai, China*

Global wild animal resources are declining due to various pressures, which will greatly affect local biodiversity and ecosystem services. Understanding local people's attitudes towards wild animal conservation in high biodiversity areas is of major importance for conservation efforts. Sampling and questionnaire survey methods were employed to examine people's attitudes toward wild animal conservation in a comparative case study of two minority ethnic groups, the Mosuo ( $n = 91$ ) and Yi ( $n = 169$ ) in Ninglang, China. Attributes influencing people's attitudes ( $p < 0.05$ ) include ethnicity, awareness, culture, level of education, religious beliefs and gender. The Mosuo are more positive towards wild animal conservation than the Yi. Women are



# Key features - data extraction

- Rather than requiring a user to read an entire paper for the thousands of potential papers that moved through System 1 screening, a user only needs to screen 35 papers
- After screening 35 papers, sentences will be surfaced toward data extraction with a “confidence” in the relevance for the user to accept or reject that data extraction tag

## ◀ Data Extraction - Label Review

Opinions on legality principles considered in the FLEGT/VPA policy in Ghana and Indonesia

LABEL SUMMARY

Sentences related to outcome\_type: subjective well-being Under Review

**1. Confidence: 96%**

by governments, is to establish an MPA and then invite local stakeholder groups to provide inputs on its management [ 16 ] . The advantages of a participatory design process include greater ownership of the results by stakeholders , increased compliance with resource regulations , and greater flexibility in resource management [ 15 ] , but without local knowledge of an MPA 's function and likely impacts , and without positive attitudes towards an MPA , participation is unlikely to be fruitful or sustained [ 12 ] . To build knowledge and positive attitudes about marine conservation , one approach is to invest in education and outreach linked to specific conservation

ACCEPT

SKIP

REJECT

**2. Confidence: 83%**

The advantages of a participatory design process include greater ownership of the results by stakeholders , increased compliance with resource regulations , and greater flexibility in resource management [ 15 ] , but without local knowledge of an MPA 's function and likely impacts , and without positive attitudes towards an MPA , participation is unlikely to be fruitful or sustained [ 12 ] . To build knowledge and positive attitudes about marine conservation , one approach is to invest in education and outreach linked to specific conservation strategies , such as MPA design and management . A number of conservation organizations are known to use this approach in marine

**3. Confidence: 81%**

The authors are grateful for the students from Universitas Papua in Manokwari , and Universitas Kristen Papua and Al Amin in Sorong who conducted the interviews , and to previous outreach staff who supported them , including A.S. Hadi , M.R. Saleh and M. Korebima . Thanks also go to Nixon Watem from Kofiau for helping with surveys and Leah Samberg for helpful advice and edits . The education and outreach activities , the perception monitoring surveys , and the analysis of the survey data were funded by NOAA , USAID , the Walton Family Foundation , the Schooner Foundation , and the HRH Foundation 's Harry and Shirley Hagey . The funders had no role in the study design ,

Wrapping it up

# Key features - data export

A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R
study_id	deduplicat	citation_scre	fulltext_scre	data_extract	data_source	data_source	data_source	citation_title	citation_abst	citation_auth	citation_jour	citation_jour	citation_pub	citation_key	fulltext_file	fulltext_excl	Intervention
144099	not_duplicate	included	not_screened	not_started	database	McKinnon_Ir	http://www.iz%a	review (this paper ad	ChamNA, ?N	unknown			2000	AntifoulingBallast watersBiofoulingCrassostrea gigas			
144100	not_duplicate	included	not_screened	not_started	database	McKinnon_Ir	http://www.iz%economic fishery resou	ReeNA, ?NA	unknown				2003	Carrying capacityDissipative structureEcological crisis			
144101	not_duplicate	included	not_screened	not_started	database	McKinnon_Ir	http://www.iz%revampin the decline ii	Mark, R.; Mu	unknown				2012	Community-based conservationprotected areasCree			
144102	not_duplicate	included	not_screened	not_started	database	McKinnon_Ir	http://www.iz%linking hu this paper ex	Broitman, B.	unknown				2013	ecosystem servicesnatural resource managementeco			
144103	not_duplicate	included	not_screened	not_started	database	McKinnon_Ir	http://www.iz%quantifyii a global ocea	Lee, J.; Rogei	unknown				2010	GIS-MCAGIS-based multi-criteria analysisFuzzy setsHu			
144104	not_duplicate	included	not_screened	not_started	database	McKinnon_Ir	http://www.iz%navigatin recreational	PolleNA, ?NA	unknown				2005	wildlifedzoosnoseseemerging diseasedisease ecologyHU			
144105	not_duplicate	included	not_screened	not_started	database	McKinnon_Ir	http://www.iz%monitorir although a ci	Bricker, S. B.	unknown				2007	iz%2000/60/ECCoastal watersE.U. Water Framework			
144106	not_duplicate	included	not_screened	not_started	database	McKinnon_Ir	http://www.iz%planning rapid ecologi	CabanbaNA, unknown					2007	iz%A priori biophysical assessmentSabahSmall-scale r			
144107	not_duplicate	included	not_screened	not_started	database	McKinnon_Ir	http://www.iz%evaluatio this chapter	BeekJ, F. A. f	unknown				2005	iz%ABCAallowable biological catchEvaluationFisheries			
144108	not_duplicate	included	not_screened	not_started	database	McKinnon_Ir	http://www.iz%applicatio the pre-aral	AssenovNA, unknown					2012	iz%Accommodation sectorAssessmentCorporate soci			
144109	not_duplicate	included	not_screened	not_started	database	McKinnon_Ir	http://www.iz%sensuring the term 'gei	Azam, I.; Firri	unknown				2012	iz%Aceh GreenCLUZCommunity-basedConservation p			
						McKinnon_Ir	http://www.iz%linking co the marine a	Campbell-Sm	unknown				2012	iz%AcehCommunity-based conservationCoral reefFis			
						nnon_Ir	http://www.iz%two solitu rapidly grow	HyndmaNA, unknown					2010	iz%AcehConflictDisastersHumanitarian assistanceInte			
						nnon_Ir	http://www.iz%facilitatin marine habit	Amezaga, J. I	unknown				2012	iz%Action researchAdaptive cycleAdaptive managem			
						nnon_Ir	http://www.iz%foraging e a cornerston	BriggNA, D. F	unknown				2007	iz%Activity patternDivingForaging behaviourGeolocat			
						nnon_Ir	http://www.iz%policy opi worldwide in	BoeNA, ?NA	unknown				2010	iz%Adaptation policiesClimate changeRiversWetland			
						nnon_Ir	http://www.iz%marine m we look at th	ArmitagNA, I	unknown				2011	iz%AdaptationAdaptive co-managementCollaboratio			
						nnon_Ir	http://www.iz%health im this introduc	HoughtoNA, unknown					2011	iz%AdaptationAdaptive reuseClimate changeEvidenc			
						nnon_Ir	http://www.iz%life on th the fast-grow	Dixon, ?NAR	unknown				2003	iz%AdaptationAfricaAgricultureForestsGlobal climate			
						nnon_Ir	http://www.iz%precursor by using an i	Berkelmans, unknown					2005	iz%AdaptationBelief networkCommunity composition			

## Export

<b>Total studies</b> .....	5799
database .....	5799
<b>Unique studies</b> .....	5765
<b>Screened citations</b> .....	3799
Excluded citations .....	2068
<b>Screened fulltexts</b> .....	73
Excluded fulltexts .....	10
<b>Studies with data extracted</b> .....	32

### Exclusions by reason

Useful .....	10
--------------	----

REFRESH

CSV

# What can you do with your data export?



<http://natureandpeopleevidence.org/ebc-dataportal-current/>

# User Engagement

- Share their experiences using Colandr
- Partner on measurement and evaluation studies
- Develop the next phase of evidence based decision-making tools.



<https://www.colandrcommunity.com/research-partners.html>

# Research protocols currently happening

- We're current testing colandr a number of different ways and we'd love to add you to our studies!
- Protocols are testing colandr against “standard” review processes
  - University of Illinois work ([presentation at CEE, and we are recruiting for another!](#))
- Protocols are looking at pain points of using colandr
  - User studies under development
- Many more ideas!

# Join our community!

**COLANDR  
COMMUNITY**

Code

Manuals + guidance

Engagement

Updates

Converse

Any issues, email  
[colandrteam@gmail.com](mailto:colandrteam@gmail.com)

Note we are volunteers and  
cannot solve every  
problem, but we try!

[www.colandrcommunity.com](http://www.colandrcommunity.com)

New research and data is being generated at an exponentially increasing rate - how can scientists, practitioners, and policy-makers rapidly and efficiently comb through this information to find what they need to make more-informed decisions?

***Colandr is an open access machine-learning assisted app for conducting systematic syntheses of evidence from primary and grey literature sources***

## SMART SORTING & TEXT MINING

Colandr has a sophisticated back-end that uses a state-of-the-art data science (natural language processing and GloVe vectors)

## OPEN ACCESS & OPEN SOURCE

Colandr is freely accessible and provides a collaborative community for users and developers

## NO BLACK BOX PROCESSES

Colandr helps speed up review process but lets the user have final say over what is ultimately included

*Colandr is supported by the Science for Nature and People Partnership (SNAPP), DataKind, and Conservation International*

*For questions, please contact:  
Samantha Cheng ([cheng@nceas.ucsb.edu](mailto:cheng@nceas.ucsb.edu)) (SNAPP)  
Caitlin Augustin ([caitlin@datakind.org](mailto:caitlin@datakind.org)) (DataKind)*

