



Identifying Topics and Trends in the Study of Common-Pool Resources Using Natural Language Processing

JOSHUA LAMBERT 

GRAHAM EPSTEIN 

JENNIFER JOEL 

JACOPO BAGGIO 

**Author affiliations can be found in the back matter of this article*

RESEARCH ARTICLE

]u[ubiquity press

ABSTRACT

The rapid growth of the literature on the commons poses an immense challenge for the synthesis and advancement of knowledge. While it may have been reasonable for previous generations of scholars to keep up to date with a literature adding thirty to fifty papers each year, there are now hundreds of papers on the commons published each year in addition to those that might be relevant to researchers on the basis of particular sectors, methods, disciplines or theories. This paper exploits recent advances in natural language processing to identify topics and trends in the literature on the commons over the past thirty years using a dynamic topic model. The results highlight the centrality of key themes concerning resources, property rights and local management, alongside growing interest in the topics of conservation and local management. The results also demonstrate the diversity of the field with topics ranging from forests, fisheries and land to urban areas and software. Overall the dynamic topic model appears to provide a useful approach for synthesizing high-level features of the literature.

CORRESPONDING AUTHOR:

Joshua Lambert

University of Central Florida,
USA

joshua.lambert@ucf.edu

KEYWORDS:

NLP; common-pool resources;
dynamic topic modeling;
resources; property rights; local
management

TO CITE THIS ARTICLE:

Lambert, J., Epstein, G., Joel, J.,
& Baggio, J. (2021). Identifying
Topics and Trends in the Study
of Common-Pool Resources
Using Natural Language
Processing. *International
Journal of the Commons*, 15(1),
pp. 206–217. DOI: [https://doi.
org/10.5334/ijc.1078](https://doi.org/10.5334/ijc.1078)

INTRODUCTION

The literature on the commons, like most academic literatures, has grown rapidly over the past decade, posing an immense challenge for synthesizing knowledge to advance the theory and practice of common-pool resource (CPR) management. In fact, a recent study published to commemorate the thirtieth anniversary of *Governing the Commons* (Ostrom 1990) found a roughly six-fold increase in the number of papers published on the commons each year between 2002 and 2019 from about 50 to over 300 per year (van Laerhoven et al. 2020). As an interdisciplinary field of research this likely represents a fraction of the publications that might be relevant to an individual scholar, likely precluding most efforts to critically read and synthesize current research. Strategies to cope with this challenge such as focusing on the publications of established researchers, may suppress innovation by neglecting insights from traditionally underrepresented groups (Hofstra et al. 2020). This paper seeks to examine opportunities to overcome this challenge by exploiting advances in natural language processing and machine learning (Blei et al. 2003, Blei 2012, Dieng et al. 2019) to uncover patterns concerning the status and development of the literature on common-pool resources.

The literature on common-pool resources, like many scientific literatures, is facing an emerging paradox involving rapid growth in the availability of information on the governance of common-pool resources (van Laerhoven et al. 2020) coupled with the relatively slow pace of advances in theory and practice (Cumming et al. 2020). For example, although more than two hundred papers related to the commons have been published each year since 2010 (van Laerhoven et al. 2020), empirical tests of individual design principles have been limited to a maximum of 63 cases (Cox et al. 2010) and 27 when extended to analyze cases in which all design principles have been coded (Baggio et al. 2016, Barnett et al. 2016, Ratajczyk et al. 2016). These numbers pale in comparison to the number of case studies that have likely been published in peer-reviewed and grey literatures since the 1990's. However, efforts to systematically code and analyze cases invariably runs into practical constraints in terms of the time and resources available to ensure the consistent coding of large numbers of cases. Natural language processing and machine learning offer a powerful approach for overcoming this challenge by providing a systematic and consistent approach for extracting, processing and analyzing text to uncover patterns and classify text using supervised and unsupervised approaches (Lin and He 2009, Rosenthal et al. 2017, Dieng et al. 2019). For example, Cheng et al. (2018) demonstrate how machine learning can be used to

automate the selection of studies for systematic reviews, while Callaghan et al. (2020) employed topic modelling to explore the potential gaps in the literature on climate change.

Natural language processing is an applied field of linguistics and artificial intelligence that is used to uncover patterns and relationships in unstructured text data (Hirschberg and Manning 2015). The field encompasses a range of approaches that can be used to extract and process text to develop tokens or features of a document, as well as several supervised and unsupervised methods in machine learning for analyzing those features. Features are mathematical representations of text within a document and corpus. A corpus is an ensemble of documents in text format. Once features are coded this information can be used to develop understanding of the characteristics of a document and the wider corpus. Unsupervised approaches, which include the approach presented here use these features to independently identify characteristics of a corpus. In contrast supervised methods begin with a labelled data set which “accurately” encodes one or more features of a document and seeks to develop a predictive model for coding unseen documents. In what follows this paper applies dynamic topic modelling, as an unsupervised method, to identify topics and trends in the literature on the commons between 1990 and 2019. The first section presents basic details about the corpus of abstracts, the development of the literature over time, and the journals in which papers are frequently published. The second section, meanwhile, reports the results of the dynamic topic model which provides details about the topics that are most prevalent based upon abstracts, as well as their evolution over time. Further details about the data and methods can be found following the discussion and in the supplementary material.

THE LITERATURE ON THE COMMONS

The literature on the commons continues to develop rapidly with a more than ten-fold increase in the number of papers published each year (*Figure 1*). In fact, the number of papers published in the last eight years (2012–2019) are roughly equivalent to the number of papers published in the previous twenty-two years. This finding is consistent with previous bibliographic studies of the commons (van Laerhoven and Ostrom 2007, van Laerhoven et al. 2020), albeit using a modified search string to gather the literature about the commons (see *Table 1*). The *International Journal of the Commons* is the top journal for publishing research on the commons with a total of more than 140 relevant papers published since its inception in 2007 and

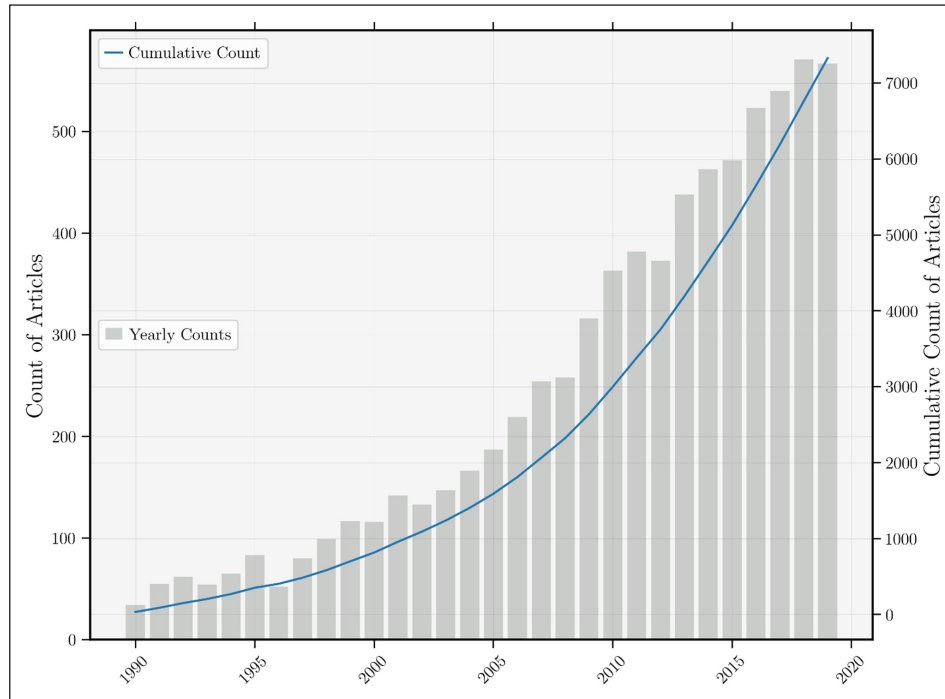


Figure 1 Annual and cumulative number of papers published in the literature on the commons. Note these figures exclude records that lacked an abstract.

JOURNAL	TOPIC 1	TOPIC 2	TOPIC 3
<i>International Journal of the Commons</i>	Resources	Land Resources	Local management
<i>Society & Natural Resources</i>	Local management	Land Resources	Conservation
<i>Ecology and Society</i>	Resources	Local Management	Land Resources
<i>Ecological Economics</i>	Resources	Games	Community Resource Management
<i>Human Ecology</i>	Land Resources	Resources	Local management
<i>World Development</i>	Resources	Local management	Land resources
<i>Marine Policy</i>	Fisheries	Local management	Resources
<i>Environmental Conservation</i>	Conservation	Local management	Species/Population Management
<i>Land Use Policy</i>	Land resources	Local management	Common property
<i>Journal of Environmental Management</i>	Local management	Resources	Conservation
<i>Environmental management</i>	Local management	Conservation	Resources
<i>Journal of Environmental Economics and Management</i>	Resources	Games	Groundwater Economics
<i>Sustainability</i>	Resources	Local management	Land Resources
<i>Oryx</i>	Species/Population management	Conservation	Local management
<i>Journal of Sustainable Forestry</i>	Local management	Conservation	Community resource management

Table 1 Distribution of topics across journals.

2019 (Figure 2). Nonetheless, more than 100 papers have been published in *Society and Natural Resources*, *Ecology and Society* and *Ecological Economics*. The relative ranking of journals is somewhat different when compared to van

Laerhoven et al. (2020), although 12 of the 15 journals listed also appear among their list of journals with more than ten papers on the commons. The omissions include the *Journal of Environmental Economics and Management*,

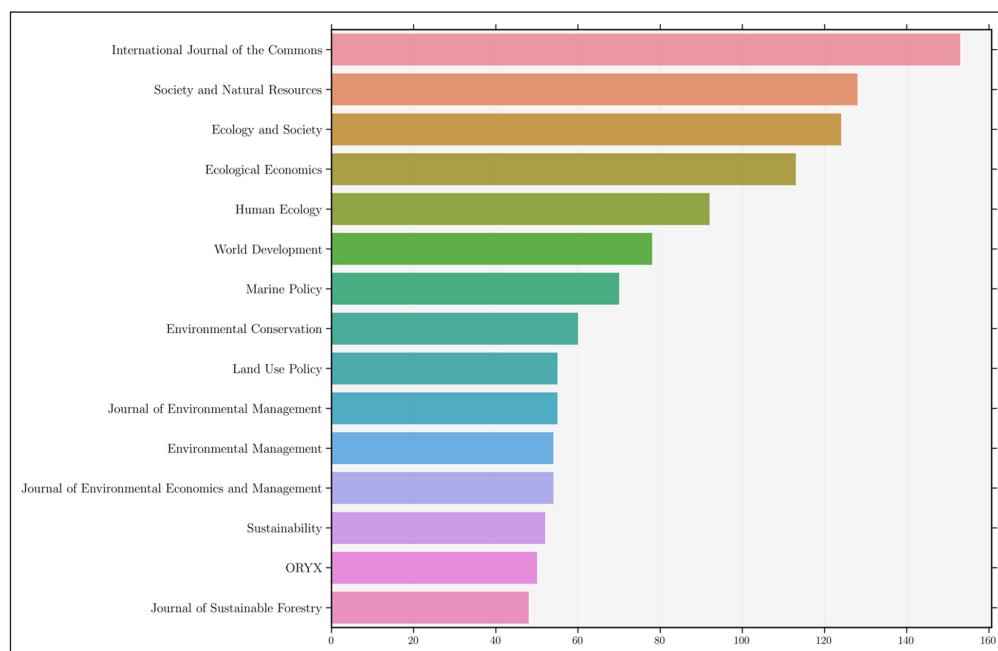


Figure 2 Number of papers published in the top 15 journals per number of publications on the commons.

ORYX and *Journal of Sustainable Forestry*, each of which was found to have more than 40 papers in our analysis. These differences are likely reflective of the expanded search string employed in this study which included several variations of community-based management across different sectors (*Table 1*), highlighting the importance of search terms used when attempting to synthesize knowledge within one or multiple fields of study.

TOPIC MODEL

To assess the evolution of the commons literature we analyze topics throughout the 30 year period analyzed (see data and methods). The results of the dynamic topic model are summarized in *Figures 3* and *4*, while further details about all 25 topics and the words associated with each topic in each three-year interval can be found in *Table S4*. The plots reveal the evolution of topics across ten time periods spanning the period 1990 to 2019 in three-year intervals. As can be seen the top five topics present variations on the general theme of common-pool resources and the management thereof. The topic of resources is unsurprisingly the most prevalent among all topics, and also contributes to other highly prevalent topics that refer to specific types of resources based upon the sector (i.e. fisheries, land and livestock – including rangelands and pastures-, forestry etc.) or specific references to the attributes of the resource itself (i.e. common resources). Common property and local management are also highly prevalent across the corpus of abstracts which speaks to rights that groups hold with respect to the use and management of resources and the scale at which

management processes take place, respectively.

Although common property and land and livestock commons feature among the most prevalent topics, they are also among the topics that have experienced the greatest declines in terms of prevalence. Common property as a topic has declined fairly consistently since the 1990's when it accounted for approximately 12% of all topics to about 8% in the current time period. Land and livestock commons, land resources and groundwater economics have also experienced declines of more than two percentage points since the early 1990's. In contrast, the topic of conservation has experienced the largest gains, increasing from about 1.8% of all topics between 1990–1992 to over 5% in the 2017–2019 time slice. Interestingly local management, which featured among the top five topics overall has also grown in prevalence by approximately 3% since 1990. When combined with the declines in common property and land and livestock commons, this suggests that the literature appears to have given greater priority to the scale of management over the attributes of resources (land and livestock commons, land resources) and property rights (common property) over time.

Finally, *Table 1* outlines how topics are distributed across the top 15 journals based upon the average prevalence of topics within papers published in the journal. Unsurprisingly the topics generally reflect the prevalence of topics across the corpus and the explicit focus of the journal. For example, the topic of fisheries is highly prevalent in the journal *Marine Policy*, while conservation is prevalent in *Environmental Conservation*. Resources and local management is the highest ranked topics for five and four of the top fifteen

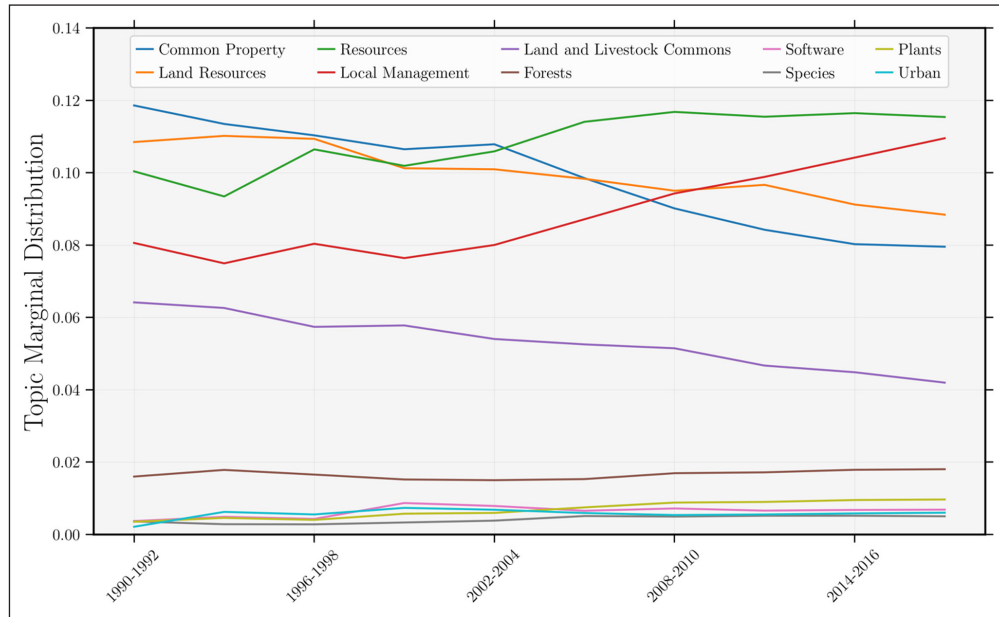


Figure 3 Marginal distribution of topics over time for the top 5 and bottom 5 topics identified in the topic model. The top five and bottom five topics are listed on the first and second row of the legend, respectively.

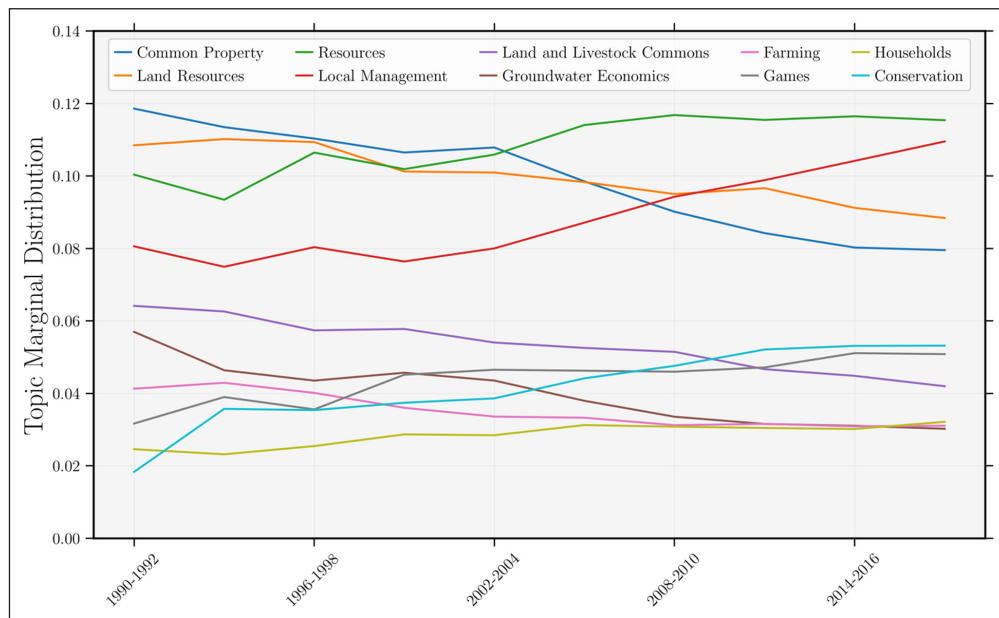


Figure 4 Marginal distribution of the top 5 topics with increasing and decreasing prevalence over time. Topics with increasing and decreasing prevalence are listed on the first and second row of the legend, respectively.

journals, respectively. The topic of games is particularly prevalent in *Ecological Economics* and the *Journal of Environmental Economics and Management*; the latter of which also appeared to include a relatively large number of papers focused on the topic of groundwater economics. Finally, the management of species and populations was particularly prevalent in conservation-oriented journals, while somewhat surprisingly the *Journal of Sustainable Forestry* appears to focus more on management and conservation than it does forests.

DISCUSSION

This paper which applied a systematic approach to examine the evolution of topics in the literature on the commons over the past thirty years using natural language processing and a dynamic topic model represents one of most extensive efforts to synthesize this literature. The results identify the top 25 topics found within the literature and indicates trends in the distribution of those topics over time and within specific journals. Furthermore this study

complements other manually supervised bibliometric studies on the literature on the commons (van Laerhoven and Ostrom 2007, van Laerhoven et al. 2020) that benefit from the supervision of experts in guiding the review process towards specific topics of interest, but that are also subject to the inherent biases of researchers in specifying those topics and potential limitations in their ability to identify latent topics or themes. Dynamic topic modelling, on the other hand, helps to address some of these issues, but also involves a number of limitations itself which are further detailed below.

First, although the dynamic topic model provides tools for identifying and determining the number of topics within a corpus of text, it does not provide an objective approach for labelling those topics. Instead, topic labels need to be assigned by the authors based upon the ensemble of words associated with each topic and their associated ranking and relevance score. This process was straightforward for topics such as forests, where the term forest consistently ranked highest. However, it was far more challenging for the topic labelled subsystems, where the term itself does not appear and instead consisted of a variety of terms related to the social, economic, environmental and political aspects of common-pool resource problems. Second, we removed a number of common and user-defined stop words prior to running the dynamic topic model, the latter of which may have a significant impact on the number and nature of topics identified. It is important to note that word inclusion/exclusion criteria may affect the results. Choices need to be documented and theoretically or purposely grounded. For example, we considered removing many of the words that formed part of our search string, such as “common-pool resource”, “commons” and “community-based”, as they were likely to emerge as a topic given the search parameters and potentially lead us to overlook other topics. However, we ultimately elected to retain these terms to allow for the possibility of examining the prevalence of different framings over time.

Third, as is evident in [Figure 1](#), the peer-reviewed literature on the commons, has grown rapidly in recent years and thus topics that emerge from the model are more likely to reflect topics that have been more relevant in recent time slices. Fourth, there is no single agreed upon definition that clearly distinguishes the literature on the commons from other literatures. The search string developed in this paper to identify this literature is a reflection of previous studies (van Laerhoven et al. 2020), a small expert survey (Table S2) and the expertise of the research team. However, as was noted by one reviewer we failed to include terms related to community based rangelands, pastures or grazing systems. Although it appears that the addition of these terms would have a minimal impact on the results as it would add fewer

than 30 papers to the corpus, it nonetheless reiterates the point that the results reflect topics found in abstracts using particular search strings and may not necessarily apply to different conceptualizations of the “literature on the commons”. Furthermore, it is worth noting that scholarly research inherently contains some bias given the prevalence of statistically significant results. Finally, topic modelling is sensitive to specific parameters employed, for example, the number of topics and the chain variance – a parameter that increases/decreases the variance of words for each topic modelled over time – which effects the evolution of topics over time can affect the results presented here.

Notwithstanding the limitations discussed above, this paper provides several insights into the topics and trends that have emerged in the literature on the commons over the past thirty years. First and perhaps foremost, the dynamic topic model shows how the literature has evolved from an initial focus on basic principles related to the nature of common-pool resources and common property systems to increasingly adopt a more problem-orientation to examine issues in conservation and local management. For example, in 1992 Schlager and Ostrom (1992), developed a highly used conceptual framework to clarify the bundle of rights associated with common property regimes and distinguish them from open-access and private property regimes. By contrast, in 2019 Edella Schlager collaborated with several scholars to examine problems related to the provision of monitoring and enforcement functions in polycentric watershed management (Hanlon et al. 2019). Although this clearly neglects Dr. Schlager’s earlier empirical research (Schlager et al. 1994) and more recent contributions to theoretical development (DeCaro et al. 2017) it nonetheless reflects the general evolution of the literature.

Second, the analysis reveals considerable diversity in terms of how scholars organize their research on common-pool resources. As discussed each of the top five topics represent slight variations on the general overarching themes of resources (resource, land resources and land and livestock resources) or institutional arrangements (common property and local management); the implications of which are somewhat mixed. On the one hand, it suggests that concerns about conflating common-pool resources with the institutions that are used to govern them (Vaccaro and Beltran 2019) might be overstated, at least within the literature examined. The fact that the topic model distinguishes between scale invariant property rights regimes and scale-dependent management systems might indicate that scholars have been careful to differentiate between the two. On the other hand, they may also be indicative of fragmentation in which knowledge and theory

is developing separately due to the lack of a common language (Ostrom 2009, Poteete et al. 2010).

Third, several interesting or unexpected topics emerged from the dynamic topic model. The topic labeled subsystems, for example, is on average the seventh most prevalent topic across the time slices, with relatively little change over time. The topic itself, consists of words such as environmental, social and political which speak to the diversity of subsystems that jointly influence how human actors interact with the environment and each other. The relative stability of the concept is somewhat surprising given that the SES framework, which explicitly identifies several subsystems, was not published until 2007 (Ostrom 2007). However, complexity and contingency have been longstanding features of both Elinor and Vincent Ostrom's work on the commons and polycentric governance (Sproule-Jones 2005, McGinnis and Ostrom 2012, Frischmann 2013) and it would seem that the SES framework has had little impact on its prevalence in this literature. Interestingly, the concept of "the commons" emerged as a distinct topic apart from those related to resources and governance. Instead, the topic of "the commons" appears to be used to refer either to changes in dominant property rights regimes (Bush and Sabri 2000, Anwar 2012) and/or offered as an alternative to privatization and state control (Hanna 1990, Bazzoli 2018). Finally, it is worth noting that despite longstanding concerns regarding the focus on small-scale systems (Young 2002, Blaikie 2006), that global environmental issues have been a small, but non-trivial part of the literature on the commons since the 1990's.

The topic model does, however, reaffirm some concerns that have been expressed concerning the lack of historical perspectives (Johnson 2004), attention to issues of power and inequality (Clement 2010) and general neglect of the mechanisms and processes that underlie sustainable common-pool resource management. Although there have been several attempts to address these gaps, including long-term studies of the development and decline of common-property governance (De Moor 2008, 2015), theoretical and empirical examinations of the relationship between power and common property governance (Epstein et al. 2014, Kashwan 2015, 2016, Bennett et al. 2018), and development and application of approaches for examining governance dynamics across networks of action situations (McGinnis 2011, McCord et al. 2017, Cole et al. 2019, Epstein et al. 2020); none of these appeared among the top 25 topics in the literature.

Fourth and finally, there are several potentially important lessons to be learned by comparing the results of this unsupervised analysis to a recent review of words contained in abstracts in the literature on common-pool resources (van Laerhoven et al. 2020). On the one hand,

unsupervised methods in machine learning provide an independent method for validating or triangulating findings across different methods, an issue that has frequently been highlighted as important in the literature on the commons (Poteete et al. 2010). For example, van Laerhoven et al. (2020) analyzed the distribution of publications across the big five sectors, which they identify in order of prevalence as water management, fisheries, forestry, rangeland and irrigation. The dynamic topic model, meanwhile, revealed a number of similar topics including forests, fisheries, water, farming, land resources and groundwater economics. Although the topics themselves are generally similar, there are important differences in terms of their construction and ranking. For instance van Laerhoven et al. (2020) distinguish between irrigation and water management; while the topic model appears to combine these into a single topic, which is distinguished from another topic on groundwater economics. Similarly, the topic model indicates that land resources as opposed to water management is by a fairly decent margin the most prevalent sector, even if we were to combine groundwater economics and water into a single topic. They identify several types of new commons or sectors that appear at least partially related to topics found in the topic model, including urban (urban), climate change (global), biodiversity (conservation), and digital commons (software).

It is important to reiterate that the dynamic topic model presented here simply uncovers patterns in a corpus of text, but lacks guidance or tools for determining the theoretical or practical importance of those topics. For instance, while van Laerhoven et al. (2020) examine the prevalence of different types of environmental governance systems (i.e. markets, state and communities) in the literature, and the extent to which different design principles are discussed, the dynamic topic model fails to provide corresponding insights. Instead the topic model identifies several topics that present slight variations on the theme of resources and management/property rights. While these may be indicative of a fragmented literature, or the need for a nuanced understanding of resources and corresponding governance systems; they nonetheless lack a certain theoretical substance when compared to an analysis of the design principles. Supervised methods, in which human agents generate specific dictionaries (similarly to a codebook – (Ratajick 2016) and use these to train algorithms to relate text features to concepts of theoretical and/or practical importance may help to address some of these shortcomings. Once again, natural language processing as all machine learning techniques heavily depend on their implementation and training. Dynamic topic modeling offer powerful tools for synthesizing knowledge on the commons, but are wholly dependent

upon the careful and detailed work of researchers developing the text that inform them.

DATA AND METHODS

The data used in this study was extracted from Scopus and the Web of Science Core collections using the search strings presented in [Table 2](#). These databases were chosen over alternatives, such as Google Scholar, based upon previous studies indicating they provide higher precision, recall and reproducibility (Gusenbauer and Haddaway 2020). The Scopus search returned a total of 7,111 records, while the Web of Science search returned 5,520 records. The data was subsequently merged into a single record of abstracts between 1990 and 2019. The search retrieved a few papers with a publication date of 2020 as they were accepted and published first online in 2019, we disregarded these articles. We then proceeded to remove duplicate records that were found in both databases by first removing records with identical titles and/or abstracts, and then analyzing similarities across articles. In order to assess whether abstracts were similar we first utilized a matching method based on how similar the titles were. This allowed for a relative score of pairwise consistency that alleviated some grammatical inconsistency between the source databases. Secondly, we used a term frequency inverse distance frequency method with cosine similarities to determine how similar abstract texts were. Records with a similarity score between 80 and 95 were reviewed manually to determine a cut-off value. The results of this review are reported in Figure S1, which was used to select a cut-off value of 87. After this process a total of 7,331 unique records remain.

Next, we pre-processed the data in order to facilitate analysis of topics and trends in the literature on common-pool resources. First we used standard pre-processing tools to remove common English stopwords (such as the, an, etc.), transform all letters to lower case, and eliminate all non-alphabetic words using Gensim and NLTK packages in python (Loper and Bird, 2002; Rehurek and Sojka, 2010). We also added a specific set of user-defined stopwords (see

Table S1) that we deemed less relevant to defining topics related to common-pool resources. For instance, words such as research, hypothesis and copyright were removed as they were expected to be common and/or relatively uninformative. We also considered removing words such as commons and common-pool resources, as they were part of the search string. However, we ultimately chose to include them to allow for the possibility of comparing different framings of the literature (i.e. the commons vs. community-based management). Finally we lemmatized words as opposed to stemming using the lemmatization modules in the NLTK Python package “en_core_web_lg” (Bird et al. 2009). Stemming is a fairly straightforward process of reducing words to their base or root form using predetermined rules (Porter 2001), often by removing suffixes. For example, depending upon the specific implementation both harvesting, and harvester could be stemmed to harvest. In contrast lemmatization identifies the part of speech of a word and uses that information to distinguish between words that refer to actions such as harvesting and those that refer to people, places or things such as harvester. In our study, we use lemmatized words as lemmatizing is more accurate when it comes to assessing specific topics that are more prevalent in a specific publication.

Following lemmatization of text, we proceeded to analyze the data. First, we analyzed general trends in the literature, including the number of papers published each year, the journals in which papers are published and the most common words that appear in abstracts. We then continued to develop dynamic topics models by first determining the optimal number of topics from the set of 20, 25 and 30 topics on the basis of topic coherence using the intrinsic UMass measure (Mimno et al. 2011). Topic coherence allows us to assess individual topics by assessing their degree of semantic similarity. That is, topics should be “internally consistent” rather than statistical artifacts (Stevens et al. 2012). It does so by assessing document co-occurrence. In fact, the U-Mass measure counts the number of pairs of documents in which a specific word w appear together, and then divides it by the overall number of documents in which the same word appears. In a nutshell, if two words are part of the same topic, they

SCOPUS	WEB OF SCIENCE
TITLE-ABS-KEY ("common pool resourc*" OR "community based natural resourc*" OR "community based resourc* manag*" OR "community based fish*" OR "community based forest*" OR "community based irrig*" OR "community based water*" OR "community based wildlife*" OR "community based conservation" OR "common property resourc*" OR "common property system" OR "common property inst*" OR {the commons AND NOT "house of" AND NOT "parliament"}) AND PUBYEAR < 2020 AND PUBYEAR > 1989	PY = 1990–2019 AND (TS="common pool resourc*" OR TS="community based natural resourc*" OR TS="community based resourc* manag*" OR TS="community based fish*" OR TS="community based forest*" OR TS="community based irrig*" OR TS="community based water*" OR TS="community based wildlife*" OR TS="community based conservation" OR TS="common property resourc*" OR TS="common property system" OR TS="common property inst*" OR TS= ("the commons" NOT "house of" NOT "parliament"))

Table 2 Research Design.

should show up “together” more often than two words that are not part of the same topic. Formally:

$$\text{UMASS Score } (w_i, w_j) = \ln \frac{D(w_i, w_j) + 1}{D(w_i)}$$

where the +1 is used to avoid $\ln(0)$. Topic coherence scores for each time period and number of topics are shown in Table S3 of the supplementary material, resulting in the selection of a model with 25 topics.

The dynamic topic model is a technique that allows for an analysis of topic evolution over time. The techniques used are based on the work of Blei and Lafferty (2006). Since we found that 25 topics allowed for the most appropriate coherence, only a few additional parameters need to be defined before the model can be implemented. First we had to decide on the time slices necessary to assess the topic modeling at. Since yearly time slices would result in an over distribution of abstracts in later years, and drastically fewer in early years, we settled on a three-year time slice. This choice helped to keep coherence scores more stable across time slices. The parameter was then set as an array of frequency counts for each period.

In addition to number of topics and time slices the final parameter for the dynamic topic model that we adjusted is the chain variance. This value represents the Gaussian distribution effecting how sensitive the model is to topical changes over time. A higher value gives the models a greater degree of variability regarding the terms which make up the topic. We altered the chain variance from the default .005 to .05 to grant a more useful depiction of commons topics over time.

Topic labels were assigned by considering the top ten words associated with each topic, and the order thereof. In a few challenging cases, the authors also considered the relevance of each term to distinguish between potentially similar topics. For instance, the words resource and common rank highly for both topic 1 and topic 5; but the relevance scores for resource are approximately 2 times higher for topic 1 and roughly equivalent for topic 5, and also include terms such as land, pasture and livestock. Thus, we assigned the labels resources and land and livestock commons to topic 1 and 5, respectively.

ADDITIONAL FILE

The additional file for this article can be found as follows:

- **Supplementary Material.** Replication files can be found here: <https://github.com/JELambert/IJC-Replication---2021>. DOI: <https://doi.org/10.5334/ijc.1078.s1>


ACKNOWLEDGEMENTS


JL and GE acknowledge support from the Preeminent Postdoctoral Program at UCF. We also would like to thank the anonymous reviewers who gave helpful and informative comments that improved this manuscript.


COMPETING INTERESTS


The authors have no competing interests to declare.

AUTHOR AFFILIATIONS

Joshua Lambert  orcid.org/0000-0002-7822-0420
University of Central Florida, USA

Graham Epstein  orcid.org/0000-0003-3431-8511
University of Central Florida, USA

Jennifer Joel  orcid.org/0000-0001-9963-4238
University of Central Florida, USA

Jacopo Baggio  orcid.org/0000-0002-9616-4143
University of Central Florida, USA

REFERENCES

- Anwar, N. H.** (2012). State power, civic participation and the urban frontier: The politics of the commons in Karachi. *Antipode*, 44: 601–620. DOI: <https://doi.org/10.1111/j.1467-8330.2011.00920.x>
- Baggio, J., Barnett, A. Perez-Ibarra, I. Brady, U. Ratajczyk, E. Rollins, N. Rubiños, C. Shin, H. Yu, D., & Aggarwal, R.** (2016). Explaining success and failure in the commons: the configural nature of Ostrom’s institutional design principles. *International Journal of the Commons*, 10. DOI: <https://doi.org/10.18352/ijc.634>
- Barnett, A., Baggio, J., Shin, H., Yu, D., Perez-Ibarra, I., Rubiños, C., Brady, U., Ratajczyk, E., Rollins, N., & Aggarwal, R.** (2016). An iterative approach to case study analysis: insights from qualitative analysis of quantitative inconsistencies. *International Journal of the Commons*, 10. DOI: <https://doi.org/10.18352/ijc.632>
- Bazzoli, N.** (2018). Struggling for housing in the changing neighbourhood. The political potential of urban commons. *ACME-AN INTERNATIONAL E-JOURNAL FOR CRITICAL GEOGRAPHIES* 17: 269–291.
- Bennett, A., Acton, L., Epstein, G., Gruby, R. L., & Nenadovic, M.** (2018). Embracing conceptual diversity to integrate power and institutional analysis: Introducing a relational typology. *International Journal of the Commons*, 12, 330–357. DOI: <https://doi.org/10.18352/ijc.819>

- Bird, S., Klein, E., & Loper, E.** (2009). *Natural language processing with Python: analyzing text with the natural language toolkit*. "O'Reilly Media, Inc".
- Blaikie, P.** (2006). Is Small Really Beautiful? Community-based Natural Resource Management in Malawi and Botswana. *World Development*, 34, 1942–1957. DOI: <https://doi.org/10.1016/j.worlddev.2005.11.023>
- Blei, D. M.** (2012). Probabilistic topic models. *Communications of the ACM*, 55, 77–84. DOI: <https://doi.org/10.1145/2133806.2133826>
- Blei, D. M., Ng, A. Y., & Jordan, M. I.** (2003). Latent dirichlet allocation. *Journal of machine Learning research*, 3, 993–1022.
- Bush, R., & Sabri, A.** (2000). Mining for fish: Privatization of the "commons" along Egypt's Northern coastline. *Middle East Report*, 20–45. DOI: <https://doi.org/10.2307/1520210>
- Callaghan, M. W., Minx, J. C., & Forster, P. M.** (2020). A topography of climate change research. *Nature Climate Change*, 10, 118–123. DOI: <https://doi.org/10.1038/s41558-019-0684-5>
- Cheng, S., Augustin, C., Bethel, A., Gill, D. Anzaroot, S., Brun, J., DeWilde, B., Minnich, R., Garside, R., & Masuda, Y.** (2018). Using machine learning to advance synthesis and use of conservation and environmental evidence. *Conservation Biology*, 32, 762–764. DOI: <https://doi.org/10.1111/cobi.13117>
- Clement, F.** (2010). Analysing decentralised natural resource governance: proposition for a "politicised" institutional analysis and development framework. *Policy Sciences*, 43, 129–156. DOI: <https://doi.org/10.1007/s11077-009-9100-8>
- Cole, D. H., Epstein, G., & McGinnis, M. D.** (2019). Combining the IAD and SES frameworks. *International Journal of the Commons*, 13, 244–275. DOI: <https://doi.org/10.18352/ijc.864>
- Cox, M., Arnold, G., & Villamayor Tomas, S.** (2010). A Review of Design Principles for Community-Based Natural Resource Management. *Ecology and Society*, 15, 38. DOI: <https://doi.org/10.5751/ES-03704-150438>
- Cumming, G. S., Epstein, G., Anderies, J. M., Apetrei, C. I., Baggio, J., Bodin, Ö., Chawla, S., Clements, H. S., Cox, M., Egli, L., Gurney, G. G., Lubell, M., Magliocca, N., Morrison, T. H., Müller, B., Seppelt, R., Schlüter, M., Unnikrishnan, H., Villamayor-Tomas, S., & Weible, C. M.** (2020). Advancing understanding of natural resource governance: a post-Ostrom research agenda. *Current Opinion in Environmental Sustainability*, 44, 26–34. DOI: <https://doi.org/10.1016/j.cosust.2020.02.005>
- De Moor, T.** (2008). The silent revolution: A new perspective on the emergence of commons, guilds, and other forms of corporate collective action in Western Europe. *International review of social history*, 53, 179–212. DOI: <https://doi.org/10.1017/S0020859008003660>
- De Moor, T.** (2015). *The dilemma of the commoners: Understanding the use of common-pool resources in long-term perspective*. Cambridge University Press. DOI: <https://doi.org/10.1017/CBO9781139135450>
- DeCaro, D. A., Chaffin, B. C., Schlager, E., Garmestani, A. S., & Ruhl, J.** (2017). Legal and institutional foundations of adaptive environmental governance. *Ecology and society: A journal of integrative science for resilience and sustainability*, 22, 1. DOI: <https://doi.org/10.5751/ES-09036-220132>
- Dieng, A. B., Ruiz, F. J., & Blei, D. M.** (2019). The dynamic embedded topic model. *arXiv preprint*. arXiv:1907.05545.
- Epstein, G., Bennett, A., Gruby, R., Acton, L., & Nenadovic, M.** (2014). Studying Power with the Social-Ecological System Framework. In Manfredo, M. J., Vaske, J. J., Reckemmer, A., & Duke, E. A. (eds.). *Understanding Society and Natural Resources*, 111–135. Springer Netherlands. DOI: https://doi.org/10.1007/978-94-017-8959-2_6
- Epstein, G., Morrison, T. H., Lien, A., Gurney, G. G., Cole, D. H., Delaroché, M., Villamayor Tomas, S., Ban, N., & Cox, M.** (2020). Advances in understanding the evolution of institutions in complex social-ecological systems. *Current Opinion in Environmental Sustainability*, 44, 58–66. DOI: <https://doi.org/10.1016/j.cosust.2020.06.002>
- Frischmann, B. M.** (2013). Two enduring lessons from Elinor Ostrom. *Journal of institutional economics*. DOI: <https://doi.org/10.1017/S1744137413000106>
- Gusenbauer, M., & Haddaway, N. R.** (2020). Which academic search systems are suitable for systematic reviews or metaanalyses? Evaluating retrieval qualities of Google Scholar, PubMed, and 26 other resources. *Research synthesis methods*, 11(2), 181–217. DOI: <https://doi.org/10.1002/jrsm.1378>
- Hanlon, J., Olivier, T., & Schlager, E.** (2019). Suspicious Collaborators: How Governments in Polycentric Systems Monitor Behavior and Enforce Public Good Provision Rules Against One Another. *International Journal of the Commons*, 13. DOI: <https://doi.org/10.5334/ijc.924>
- Hanna, S. S.** (1990). The eighteenth century English commons: a model for ocean management. *Ocean and shoreline management*, 14, 155–172. DOI: [https://doi.org/10.1016/0951-8312\(90\)90032-D](https://doi.org/10.1016/0951-8312(90)90032-D)
- Hirschberg, J., & Manning, C. D.** (2015). Advances in natural language processing. *Science*, 349(6245), 261–266. DOI: <https://doi.org/10.1126/science.aaa8685>
- Hofstra, B., Kulkarni, V. V., Munoz-Najar Galvez, S., He, B., Jurafsky, D., & McFarland, D. A.** (2020). The Diversity–Innovation Paradox in Science. *Proceedings of the National Academy of Sciences*, 117, 9284–9291. DOI: <https://doi.org/10.1073/pnas.1915378117>
- Johnson, C.** (2004). Uncommon Ground: The 'Poverty of History' in Common Property Discourse. *Development and Change*, 35, 407–434. DOI: <https://doi.org/10.1111/j.1467-7660.2004.00359.x>
- Kashwan, P.** (2015). Forest policy, institutions, and REDD+ in

- India, Tanzania, and Mexico. *Global Environmental Politics*, 15, 95–117. DOI: https://doi.org/10.1162/GLEP_a_00313
- Kashwan, P.** (2016). Integrating power in institutional analysis: A micro-foundation perspective. *Journal of Theoretical Politics*, 28, 5–26. DOI: <https://doi.org/10.1177/0951629815586877>
- Lin, C., & He, Y.** (2009). Joint sentiment/topic model for sentiment analysis. In *Proceedings of the 18th ACM conference on Information and knowledge management*, 375–384. DOI: <https://doi.org/10.1145/1645953.1646003>
- McCord, P., Dell'Angelo, J., Baldwin, E., & Evans, T.** (2017). Polycentric Transformation in Kenyan Water Governance: A Dynamic Analysis of Institutional and SocialEcological Change. *Policy Studies Journal*, 45, 633–658. DOI: <https://doi.org/10.1111/psj.12168>
- McGinnis, M. D.** (2011). Networks of Adjacent Action Situations in Polycentric Governance. *Policy Studies Journal*, 39, 51–78. DOI: <https://doi.org/10.1111/j.1541-0072.2010.00396.x>
- McGinnis, M. D., & Ostrom, E.** (2012). Reflections on Vincent Ostrom, public administration, and polycentricity. *Public Administration Review*, 72, 15–25. DOI: <https://doi.org/10.1111/j.1540-6210.2011.02488.x>
- Mimno, D., Wallach, H., Talley, E., Leenders, M., & McCallum, A.** (2011). Optimizing semantic coherence in topic models. In *Proceedings of the 2011 Conference on Empirical Methods in Natural Language Processing*, 262–272.
- Ostrom, E.** (1990). *Governing the Commons*. New York: Cambridge University Press. DOI: <https://doi.org/10.1017/CBO9780511807763>
- Ostrom, E.** (2007). A Diagnostic Approach for Going beyond Panaceas. *Proceedings of the National Academy of Sciences of the United States of America*, 104, 15181–15187. DOI: <https://doi.org/10.1073/pnas.0702288104>
- Ostrom, E.** (2009). A General Framework for Analyzing Sustainability of Social-Ecological Systems. *Science*, 325, 419–422. DOI: <https://doi.org/10.1126/science.1172133>
- Poteete, A. R., Janssen, M. A., & Ostrom, E.** (2010). *Working Together*. Princeton, NJ: Princeton University Press. DOI: <https://doi.org/10.1515/9781400835157>
- Ratajczyk, E., Brady, U., Baggio, J., Barnett, A., Perez-Ibarra, I., Rollins, N., Rubiños, C., Shin, H., Yu, D., & Aggarwal, R.** (2016). Challenges and opportunities in coding the commons: problems, procedures, and potential solutions in large-N comparative case studies. *International Journal of the Commons*, 10. DOI: <https://doi.org/10.18352/ijc.652>
- Rosenthal, S., Farra, N., & Nakov, P.** (2017). SemEval-2017 task 4: Sentiment analysis in Twitter. In *Proceedings of the 11th international workshop on semantic evaluation (SemEval-2017)*, 502–518. DOI: <https://doi.org/10.18653/v1/S17-2088>
- Schlager, E., Blomquist, W., & Tang, S. Y.** (1994). Mobile Flows, Storage, and Self-Organized Institutions for Governing Common-Pool Resources. *Land Economics*, 70, 294–317. DOI: <https://doi.org/10.2307/3146531>
- Schlager, E., & Ostrom, E.** (1992). Property-Rights Regimes and Natural Resources: A Conceptual Analysis. *Land Economics*, 68, 249–262. DOI: <https://doi.org/10.2307/3146375>
- Sproule-Jones, M.** (2005). The Concept of Contingency and the scholarship of Elinor Ostrom on the Commons: Commentary on Tom Dietz's "the Darwinian Trope in the drama of the Commons". *Journal of Economic Behavior & Organization*, 57, 231–235. DOI: <https://doi.org/10.1016/j.jebo.2004.06.020>
- Stevens, K., Kegelmeyer, P., Andrzejewski, D., & Buttler, D.** (2012). Exploring topic coherence over many models & many topics. In *Proceedings of the 2012 Joint Conference on Empirical Methods in Natural Language Processing and Computational Natural Language Learning*, 952–961.
- Vaccaro, I., & Beltran, O.** (2019). What Do We Mean by "the Commons?" An Examination of Conceptual Blurring Over Time. *Human Ecology*, 47, 331–340. DOI: <https://doi.org/10.1007/s10745-019-00081-z>
- van Laerhoven, F., & Ostrom, E.** (2007). Traditions and Trends in the Study of the Commons. *International Journal of the Commons*, 1, 3–28. DOI: <https://doi.org/10.18352/ijc.76>
- van Laerhoven, F., Schoon, M., & Villamayor-Tomas, S.** (2020). Celebrating the 30th Anniversary of Ostrom's *Governing the Commons*: Traditions and Trends in the Study of the Commons, Revisited. *International Journal of the Commons*, 14. DOI: <https://doi.org/10.5334/ijc.1030>
- Young, O. R.** (2002). *The institutional dimensions of environmental change: Fit, interplay, and scale*. Cambridge, MA: MIT press. DOI: <https://doi.org/10.7551/mitpress/3807.001.0001>

TO CITE THIS ARTICLE:

Lambert, J., Epstein, G., Joel, J., & Baggio, J. (2021). Identifying Topics and Trends in the Study of Common-Pool Resources Using Natural Language Processing. *International Journal of the Commons*, 15(1), pp. 206–217. DOI: <https://doi.org/10.5334/ijc.1078>

Submitted: 15 September 2020 Accepted: 09 April 2021 Published: 16 June 2021

COPYRIGHT:

© 2021 The Author(s). This is an open-access article distributed under the terms of the Creative Commons Attribution 4.0 International License (CC-BY 4.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited. See <http://creativecommons.org/licenses/by/4.0/>.

International Journal of the Commons is a peer-reviewed open access journal published by Ubiquity Press.

