Conservation culturomics

Richard J Ladle[1,2*†], Ricardo A Correia[1,2], Yuno Do[3], Gea-Jae Joo[3], Ana CM Malhado[2], Raphaël Proulx[4], Jean-Michel Roberge[5], and Paul Jepson[2*†]

Culturomics is an emerging field of study that seeks to understand human culture through the quantitative analysis of changes in word frequencies in large bodies of digital texts. Culturomics research can help practitioners in nature conservation respond to cultural trends, building and reinvigorating its societal relevance. We identify five areas where culturomics can be used to advance the practice and science of conservation: (1) recognizing conservation-oriented constituencies and demonstrating public interest in nature, (2) identifying conservation emblems, (3) providing new metrics and tools for near-real-time environmental monitoring and to support conservation decision making, (4) assessing the cultural impact of conservation interventions, and (5) framing conservation issues and promoting public understanding. More generally, culturomics opens up an exciting new area of research, equipping conservationists with novel tools to explore and shape human interactions with the natural world.

Front Ecol Environ 2016; 14(5): 269–275, doi:10.1002/fee.1260

Words are symbolic representations of concepts, places, or objects (Carlston 2013). Thus, the frequency with which words and phrases are used within a language provides information about their cultural saliency or visibility (see Correia *et al.* 2016), generating insights into human cultures and how they change. The study of human culture through the analysis of changes in word frequencies in large bodies of texts (termed "corpora") is known as culturomics (Michel *et al.* 2011). Research in this area has recently intensified due to the rapid spread of digital media and the digitization of a substantial proportion of the world's historical written resources.

In a nutshell:

- Culturomics is the study of human culture through the analysis of changes in word frequencies in enormous digital text databases
- Conservation researchers are exploring the potential for culturomics to provide insights into the complex interactions between human societies and the natural world
- Culturomics can be used to promote and demonstrate public interest in nature, identify new conservation emblems, provide new metrics and tools to monitor human interactions with the environment, frame conservation issues, and assess the impact of conservation interventions

[1]*Institute of Biological Sciences and Health (ICBS), Federal University of Alagoas, Maceió, Brazil;* [2]*School of Geography and the Environment, University of Oxford, Oxford, UK* [*]*(richard.ladle@ouce.ox.ac.uk);* [3]*Department of Biological Science, Pusan National University, Busan, South Korea;* [4]*Centre de Recherche sur les Interactions Bassins Versants, Université du Québec à Trois-Rivières, Trois-Rivières, Canada;* [5]*Department of Wildlife, Fish and Environmental Studies, Swedish University of Agricultural Sciences (SLU), Umeå, Sweden;* [†]*These authors contributed equally to this work*

Culturomic trends are driven by at least two major social forces: (1) *cultural change* and (2) *linguistic change*. Cultural change influences the representation of "concepts" such as "balance of nature" or "environment", and its central idea is that changes in society will influence the frequency with which words associated with public interests, fashions, issues, and values appear within representative bodies of text. At a simplistic level this may be because a concept has been discredited (eg eugenics), but more often it is the result of a complex interplay of changing demographics, social values, formal education, media representations, and so forth. Linguistic change influences the words that are used to describe concepts and may be closely linked to cultural change (eg the replacement of "chairman" with "chairperson" or "chair"), or may be culturally neutral, whereby a word is replaced in common usage by a culturally equivalent synonym (eg "biological diversity" versus "biodiversity").

Advancing conservation through culturomics: five key areas

The power of culturomics resides in its ability to provide novel insights into the functioning of society, passively crowdsourcing and locating the tone and mood of societies and other social groups. Such insights are potentially valuable to conservation as it strives to maintain its relevance in an increasingly globalized and data-rich world. Culturomics research can provide new ways for conservation managers and researchers to document (and respond to) changes in the interactions between societies and the natural environment. Here, we outline five areas where the development of the field of "conservation culturomics" could advance the practice and science of conservation.

Constituencies and public interest

Conservation-oriented constituencies must be identified in order to ensure routine access to decision makers, to promote conservation agendas, and to mobilize action to counter threats to species, sites, and ecosystems. Culturomic methods can be thought of as a form of polling, albeit one that does not involve the direct questioning of individuals. Analysis of digital corpora enables quantification of different constituencies via word frequencies, with the potential advantages – as compared with traditional polling – of greater reach, cost efficiency, and responsiveness.

A compelling illustration of this potential is Eric Phu's groundbreaking study, commissioned by the Wildlife Conservation Society, of Chinese attitudes to elephants and ivory (Phu 2014). Phu analyzed ~1.2 million social media conversations (85% of conversations were on Sina Weibo, China's equivalent of Twitter) between 1 Jan and 30 Jun 2013 containing the keywords "environmental protection", "elephants", "ivory", and "luxury". The results challenge common perceptions of China's limited conservation constituency. Unexpectedly, wildlife protection themes appeared in 18% of posts and there were 10 times as many conversations mentioning wildlife than air pollution. On the basis of statistical associations between keywords in posts, Phu found that Chinese "netizens" (net citizens) consider ivory neither as a luxury product nor as a popular item of purchase. He concluded that conservation's problem may not be how to convince 1.3 billion Chinese to stop consuming ivory, but how to mobilize the potentially massive online sentiment pool to speak out against the few that do.

More generally, culturomics can be used to track public interest in natural attributes and environmental issues. For instance, McCallum and Bury (2013) measured the saliency of 19 environment-related terms between 2001 and 2009 using Google Insights for Search (GIFS), a forerunner of Google Trends. They demonstrated a globally decreasing use of terms such as "biodiversity" and "wildlife" over the 9-year study period. Likewise, Wilde and Pope (2013) used the same methods to track interest in recreational fishing between 2004 and 2011, reporting declines in frequencies of words associated with fishing interest in nearly all of the 50 countries examined. Although the robustness of these methods and their interpretations have been questioned (Ficetola 2013), such approaches clearly illustrate the potential of culturomics to capture public interest over time frames that are sufficiently short to develop responsive conservation strategies. Moreover, culturomic techniques can quantify and link contemporary public interest in nature with trending issues in culture and politics, invigorating and widening public participation in discussions about nature and conservation.

Identifying conservation emblems

It is a well-known truism in conservation that people value what they know, and societies act to conserve heritage. Conservation flagships, icons, and emblems are an enduring strategy for mobilizing support for conservation, providing a sense of nation and place, appearing in cultural narratives and appealing to people's curiosity, wonder, and compassion (Jepson and Barua 2015).

Culturomics provides a systematic means of assessing the relative saliency of species and to identify traits associated with cultural visibility. For example, Żmihorski *et al.* (2013) ranked Polish birds and British butterflies using Google Search returns. Their models suggest that, in these taxa, internet saliency is driven by a combination of biological traits (eg body size) and familiarity. A more recent study in the US found that internet searches for the common names of 68 resident bird species were positively associated with estimates of avian population densities (Schuetz *et al.* 2015). This result strongly suggests that people's interest in a given species is being influenced by the frequency of encountering it in the wild. Studies such as these, which generate insights into the relationship between traits and internet salience, could be used to identify new or underused conservation emblems.

Culturomics also has the potential to inform conservation communication strategies. For instance, Kim *et al.* (2014) used logs of web search data (2007–2012) to assess internet saliency of 246 species classified as endangered by the South Korean authorities. The authors found very few species with high saliency – mainly associated with body size and distinct appearance. Similarly, Clements (2013) observed correlations between Google Trends data on five recently extinct animals and pledges of financial support to the World Wildlife Fund. Such investigations highlight the need to raise the profile of smaller and less "showy" species of conservation concern and the enduring power of extinction as a conservation fund-raising narrative.

Although the above examples are geographically and taxonomically limited, many of the digital corpora analyzed by culturomic tools are global in scope and can be evaluated at multiple spatial and temporal scales. Such a multi-scalar approach potentially allows researchers to map geographic similarities and discontinuities in the cultural visibility of species or other natural features, providing opportunities to identify shared natural icons and emblems. This type of information could, in turn, form the basis for international conservation partnerships, potentially leading to enhanced funding and technical support.

Environmental monitoring and valuation

Our ability to evaluate and monitor ecosystem services at large scales is constrained by limited and fragmented data, particularly in less-developed countries. However, internet practices across the world are generating huge quantities of ecologically relevant information.

Culturomic methods can capture useful data, extending existing approaches to environmental monitoring and evaluation (Galaz *et al.* 2009).

Words are routinely used to translate human–nature interactions. In a recent study, Proulx *et al.* (2014) reported the seasonal trend of internet search terms (such as *mosquitoes* and *pollen*) that convey information about the biotic environment. The authors used Google Trends to compare the number of internet queries for the word *mosquitoes* with the number of queries for *DEET* (diethyltoluamide) or *citronella* – the two active ingredients in most commercial insect repellents. Likewise, they correlated search-term results for *pollen* with those of *Zyrtec*, *Claritin*, or *Reactine* – three allergy-treating drugs sold in Canada and the US. In these examples, search terms reflected interactions between people and the biotic environment (eg mosquito bites or allergic reactions to pollen). Other text-mining tools such as Voyant™ (http://voyant-tools.org) have the potential to assess the saliency of words in online documents (eg municipal newsletters) to reveal concerns about regional environmental or cultural issues.

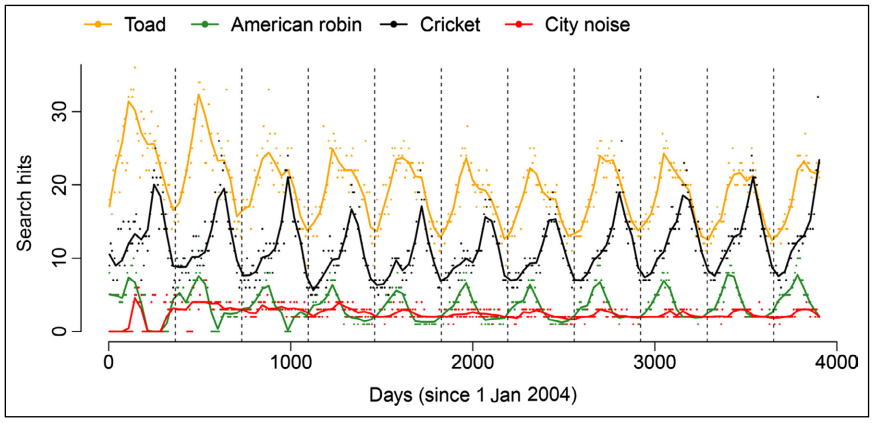

Figure 1. *Weekly time series (2004–2014) of the number of search-term hits returned by querying in Google Trends the keywords* toad, American robin, cricket, *and* city noise *for the conterminous US. Vertical dashed lines mark January 1st of each year. LOESS (locally weighted scatterplot smoothing) curves have a span of 0.05 and were applied to each time series.*

Culturomics can also potentially contribute to the emerging field of soundscape ecology (Pijanowski *et al.* 2011). By way of illustration, we quantified an interaction between urban people and their soundscape by querying Google Trends for "acoustic" search terms, such as: *toad*, *American robin*, *cricket*, and *city noise*. Results were filtered by topic and country, restricting the search to the US and associating the first three terms listed above with "amphibian", "bird", and "insect" topics, respectively. Time-series data returned clear seasonal trends for search terms associated with sound-producing animals, which was in sharp contrast with the flat (no or low slope) trend obtained for *city noise* (Figure 1). This correlative illustration suggests that urban dwellers may be attentive to seasonal changes in their soundscape despite the constant drone of city life.

Culturomics may also be appropriate for measuring cultural ecosystem services, defined as the non-material benefits people obtain from ecosystems (eg spiritual enrichment, cultural identity, recreation, aesthetic experiences, and the like) (Milcu *et al.* 2013). Long recognized as one of the key benefits that humans derive from nature, such services have proven difficult to define, measure, and integrate into existing ecosystem-service frameworks. Culturomics has the potential to provide researchers with the tools to systematically classify, quantify, and map cultural values accrued from nature at macro-scales. For example, the relative internet representation (a proxy of cultural saliency) of species can be calculated and converted into maps of cultural value based on the species assemblage at any given location (WebPanel 1).

Eventually, culturomic monitoring and assessment could be integrated into decision-support tools for land-use planners and those working in the private sector. Culturomics may be particularly effective for risk screening of extractive companies by investment markets. Several extractive companies have petroleum or mineral reserves that overlap protected areas (WWF 2015). The cultural visibility of the protected areas and/or the species that live inside their boundaries provides a proxy of the ability of environmental groups to mobilize public support. Well-supported campaigns by non-governmental organizations translate into corporate reputational risk, increased transaction costs, and even asset stranding, all of which could adversely affect investment returns. Similar tools could be developed to assess the cultural impacts of proposed development projects, thereby enhancing the "safeguarding" role of conservation.

Cultural impact assessment

In general, conservation agendas and interventions achieve greater success when they attract public interest and support. Metrics of cultural baselines and long-term impacts of conservation actions can be used to enhance the design of conservation interventions with a public dimension. Such metrics could be produced though traditional survey approaches, but these are inevitably constrained by cost, standardization of design, and difficulties of defining historical baselines. For instance, Do *et al.* (2015) investigated the cultural impact of Ramsar wetland designation in South Korea by tracking relative search volumes (RSVs) of issues related to wetland areas in the country. They found that most internet searches

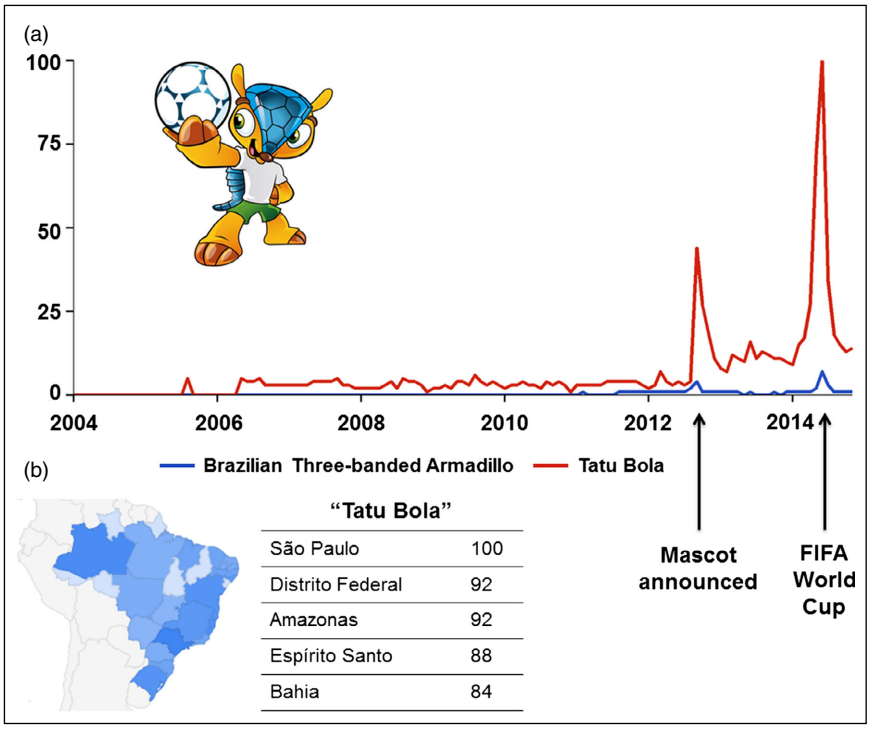

Figure 2. *Google Trends analysis of search-term data for the FIFA 2014 World Cup mascot species, the Brazilian three-banded armadillo* (tatu bola *in Portuguese). (a) Relative frequency of searches over time. (b) Relative frequency of searches for different Brazilian states. Google Trends output is returned in the form of a relative search interest based on the number of searches for the specific search term in a given period (day, week, or month) relative to the total number of searches over the specified time period.*

were related to the largest and most well-known tidal and inland wetlands, and increased substantially after their designation as Ramsar sites in 2008. Interestingly, RSVs for the largest areas were also positively associated with the number of wetland-themed articles in newspapers, which, in turn, were associated with more visitors to these wetlands. Identifying such synergies opens the prospect of a more systems-based approach to conservation policy implementation and communication.

A second example demonstrates the cultural impact of choosing a nature-based mascot for a mega-event: there was a large increase in Google searches (in English and Portuguese) for the Brazilian three-banded armadillo (*Tolypeutes tricinctus*) when it was initially unveiled as the mascot for the 2014 FIFA (Fédération Internationale de Football Association) World Cup and again during the competition (Figure 2). In an attempt by FIFA to increase environmental awareness of species conservation issues, the mascot was named Fuleco™, a combination of the Portuguese words *futebol* and *ecologia*. Protected areas of northeastern Brazil included only 8% of the three-banded armadillo distribution, and this endemic species is considered highly vulnerable to habitat loss and climate change (Melo *et al.* 2014). After lobbying by Brazilian scientists, the *Tatu-Bola* Wildlife Reserve (Refugio da Vida Silvestre Tatu Bola) was established in March 2015 to protect the armadillo.

These examples illustrate the potential of culturomics for assessing the impact of branding-related conservation actions – a designation and an emblem. The armadillo example also demonstrates the speed at which culturomics data can be collected and visualized, and the opportunities this presents for rapid response to sudden spikes in media/public interest generated by cultural or other events.

Policy and issue framing

How an issue is understood determines the solutions deemed appropriate, and how an issue is framed influences the solution's adoption by others (eg policy makers, activists). Conservation therefore needs to understand potential misconceptions and context if it is to maximize its societal impact (cf Benford and Snow 2000). By tracking the frequency of the phrases "balance of nature" and "flux of nature" using Google Search (popular representation) and Google Scholar (academic representation), Ladle and Gillson (2009) showed that, despite losing support from the wider scientific community, the balance of nature paradigm still dominated popular discourse. Such a time lag between academic and public understanding represents a potential problem for conservation, since the public may be less accepting of conservation interventions when they appear to go against deeply ingrained perceptions of conservation as a strategy to maintain ecological stability, harmony, and balance. The study also identified clear opportunities for improving conservation education by pinpointing an important misunderstanding – about the fragility/resilience of nature – to be countered.

Culturomics also has the potential to provide information on the public consequences of conservation failure. A Google Ngram analysis of five iconic North American bird extinctions (Figure 3) shows peaks in common name frequency after the presumed disappearance of some species such as the great auk (*Pinguinus impennis*). While it is unfortunate that species may be better known after they become extinct, the sustained interest in these species contributes to public support for measures aimed at avoiding future extinctions (Ladle and Jepson 2010).

Culturomics methods: challenges and limitations

Despite the exciting potential of conservation culturomics, digital social research is a relatively new field and many of the ontological and epistemological challenges are still being identified and worked through (Kitchin 2014). Culturomic methods have three main components: (1) identifying the digital corpus or corpora most appropriate for answering the research question, (2) accessing or extracting the data from these corpora, and (3) analyzing the data.

Currently, the corpora most commonly used by environmental scientists are the World Wide Web, Twitter, Google Books, and scientific databases such as Web of Science. The enormous size of many of the corpora and/or their corporate ownership means that the data are accessible only to researchers, often in pre-analyzed form, via proprietary tools (eg Google Search, Microsoft Bing) and/or application programming interfaces (APIs; software that allows application programs to interact and share data). For example, although Google Trends allows users to generate relative frequencies and geographies of search-term use, the data on which these frequencies are based are not publically available. The same is true for Google Ngram, which returns relative frequencies of words appearing in a giant corpus of digitized books published between 1500 and 2012. Although available for download, the pre-analyzed Ngram frequencies are generated by complex algorithms, and access to the underlying books is very limited.

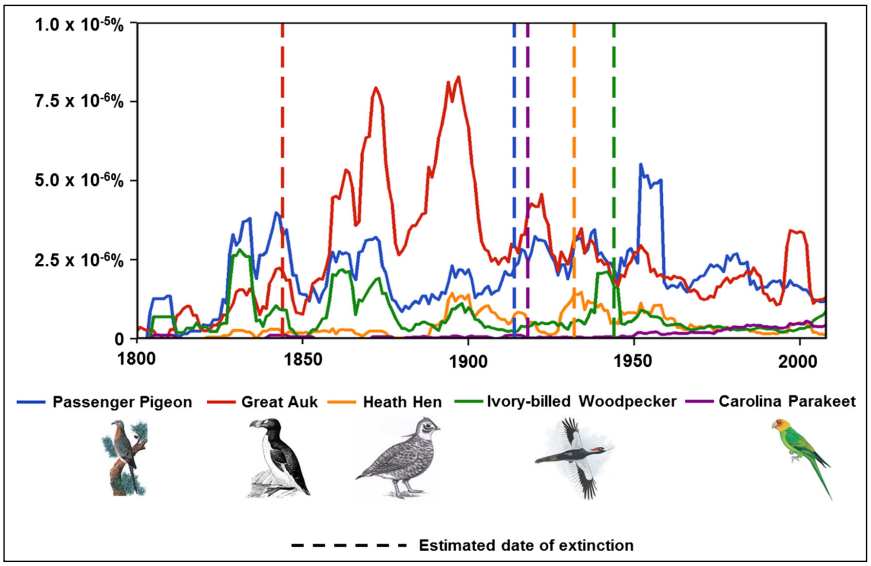

Figure 3. *NGrams of (selected) extinct North American bird species. Presumed extinction dates indicated with dashed lines. Data generated from Google's NGram Viewer, which allows users to search the enormous corpus of digitized books that stretches back to 1800.*

Even when original data are accessible, such as from the micro-blogging platform Twitter, data retrieval may be incomplete and/or unrepresentative. This is partly because the APIs of internet corporations are designed to protect the primacy of their approved data providers. Brooker *et al.* (2015), for instance, observed temporal data gaps in "tweets" returned by the Twitter API, unavoidably compromising the completeness of the downloaded dataset. Similar data gaps are generated by user privacy settings on social media platforms. Additionally, web-based search engines used to determine online visibility can bias data retrieval through personalization algorithms that return data based on user profile and location. There are various strategies for dealing with incomplete or biased data, such as deactivating personalization services or using custom search engines, but none of these strategies are completely effective in all circumstances.

Another critical challenge for culturomics research is dealing with semantic complexity. Language is inexact: it may contain multiple terms for the same thing and many words have multiple meanings and associations. We term this the "-onym" challenge. For example, the word "jaguar" is a theronym for the South American felid (*Panthera onca*) and a luxury automobile brand. In the mid-1940s, SS Cars Limited changed its name to Jaguar Cars Limited due to the former's undesirable similarity with the "SS" abbreviation associated with a Nazi German paramilitary force. The new logo was chosen on account of the perceived positive characteristics of the jaguar – including strength, speed, and beauty. Clearly, searching for "jaguar" in many of the available digital corpora would generate considerable "noise" relating to the car brand. Such noise may or may not be of interest, depending on whether the research focus is the jaguar as an ecological species or as a cultural icon.

As well as word frequency, conservation culturomics is also interested in sentiment. However, the latter is particularly difficult to automatically assess, owing to the semantic complexity of language and the frequent use of abbreviations, misnomers, misspellings, slang, and sarcasm (Murphy 2014). Although recent computational advances (eg Thelwall *et al.* 2010) enable researchers to categorize and measure the overall sentiment of large bodies of text through analysis of individual words (eg happy, hate), the sentiment constructed by the context in which the words are used (eg "I am/was happy") is currently at the frontier of natural language processing (eg Poria *et al.* 2014).

Another relevant challenge is dealing with the unreliability of profile (user contributed) or device generated (via inbuilt GPS or user's Internet Protocol [IP] address) location data. For example, user profile location data cannot be used as proxy for device location (Graham 2014).

Likewise, while all websites have an IP address that is registered in a particular country, many domain names are purchased through third-party providers, meaning that the actual location is uncertain. This is potentially problematic within conservation, given that researchers are typically very interested in location – for instance, sites where human–wildlife conflicts are happening or the geographies of public interest for particular species or sites.

Cross-cultural analysis also poses a considerable challenge because internet penetration (the proportion of the population that has access to the internet) and the ability to generate content vary between countries and regions, as well as among social and ethnic groups (Graham 2014). An influential strand of conservation seeks to recognize and give voice to rural and indigenous perspectives. Such groups may have limited internet access and lack visibility on digital corpora, placing them beyond the reach of culturomics research. In a similar vein, digital corpora are more complete for the major language groups (Funk and Rusowsky 2014). Thus, the potential of culturomics relates to the study of mass internet culture, may represent linguistic groups rather than geographical groups, and may not be applicable to all issues of conservation interest.

Conservation culturomics researchers need to understand and engage with the computational and technological processes through which datasets are constructed, along with issues of semantic complexity, internet geographies, and cultural engagements with nature and technology. This requires a "learning by doing" approach: acquiring the necessary skills and framing questions suited to the data (Brooker *et al.* 2015). Moreover, the constantly changing computational algorithms (eg Twitter APIs) limit replication, emphasizing the necessity for researchers to make their datasets – and other scientific outputs – accessible. Natural science evolved in a world where data were rare and expensive to collect. In contrast, culturomics is firmly situated in the 21st-century world of abundant, "messy" data that is driving new epistemologies. Indeed, culturomic research does not deal with "raw data" – it analyses data that are produced from the interactions between humans and the digital world.

Conclusions

If conservation and its physical manifestations (eg protected areas and protected species) are to remain relevant in the 21st century, the global conservation movement needs to continually restate the case for protecting nature. More generally, there is a growing sense that conservation needs to undergo a period of self-reflection and adaptation if it is to retain its cultural relevance in our rapidly changing world (Chan 2008). How conservation responds to the Information Revolution will be a critical part of this process of adaptation. Culturomics engages with new and powerful data resources and provides one avenue for conservation to interact with, and shape, the digitally mediated societal change that is underway.

Culturomic research is in its infancy. Nevertheless, there are an increasing number of conservation researchers throughout the world who are exploring its potential to provide insights into the complex interactions between human culture and the natural world. Culturomics has enormous potential not only to instigate discussion and experimentation but also to promote cooperation between conservationists, computer scientists, information engineers, and the digital humanities.

References

Benford RD and Snow DA. 2000. Framing processes and social movements: an overview and assessment. *Annu Rev Sociol* **26**: 611–39.

Brooker P, Barnett J, Cribbin T, *et al.* 2015. Have we even solved the first 'big data challenge'? Practical issues concerning data collection and visual representation for social media analytics. In: Snee H, Hine C, Morey Y, *et al.* (Eds). Digital methods for social science: an interdisciplinary guide to research innovation. London, UK: Palgrave MacMillan.

Carlston DE. 2013. The Oxford handbook of social cognition. Oxford, UK: Oxford University Press.

Chan K. 2008. Value and advocacy in conservation biology: crisis discipline or discipline in crisis? *Conserv Biol* **22**: 1–3.

Clements CF. 2013. Public interest in the extinction of a species may lead to an increase in donations to a large conservation charity. *Biodivers Conserv* **22**: 2695–99.

Correia RA, Jepson PR, Malhado ACM, and Ladle RJ. 2016. Familiarity breeds content: assessing bird species popularity with culturomics. *PeerJ* **4**: e1728.

Do Y, Kim JY, Lineman M, *et al.* 2015. Using internet search behavior to assess public awareness of protected wetlands. *Conserv Biol* **29**: 271–79.

Ficetola GF. 2013. Is interest toward the environment really declining? The complexity of analysing trends using internet search data. *Biodivers Conserv* **22**: 2983–88.

Funk SM and Rusowsky D. 2014. The importance of cultural knowledge and scale for analysing internet search data as a proxy for public interest toward the environment. *Biodivers Conserv* **23**: 3101–12.

Galaz V, Crona B, Daw T, *et al.* 2009. Can web crawlers revolutionize ecological monitoring? *Front Ecol Environ* **8**: 99–104.

Graham M. 2014. Internet geographies: data shadows and digital divisions of labour. In: Graham M and Dutton WH (Eds). Society and the internet: how networks of information and communication are changing our lives. Oxford, UK: Oxford University Press.

Jepson P and Barua M. 2015. A theory of flagship species action. *Conserv Soc* **13**: 95–104.

Kim JY, Do Y, Im RY, *et al.* 2014. Use of large web-based data to identify public interest and trends related to endangered species. *Biodivers Conserv* **12**: 1–24.

Kitchin R. 2014. Big data, new epistemologies and paradigm shifts. *Big Data Soc* **1**: 1–12.

Ladle RJ and Gillson L. 2009. The (im)balance of nature: a public perception time-lag? *Public Underst Sci* **18**: 218–42.

Ladle RJ and Jepson P. 2010. Origins, uses, and transformation of extinction rhetoric. *Environ Soc Adv Res* **1**: 96–115.

McCallum ML and Bury GW. 2013. Google search patterns suggest declining interest in the environment. *Biodivers Conserv* **22**: 1355–67.

Melo FP, Siqueira JA, Santos BA, *et al.* 2014. Football and biodiversity conservation: FIFA and Brazil can still hit a green goal. *Biotropica* **46**: 257–59.
Michel JB, Shen YK, Aiden AP, *et al.* 2011. Quantitative analysis of culture using millions of digitized books. *Science* **331**: 176–82.
Milcu AI, Hanspach J, Abson D, *et al.* 2013. Cultural ecosystem services: a literature review and prospects for future research. *Ecol Soc* **18**: 44.
Murphy J. 2014. Ten things survey researchers should know about Twitter. In: Woodfield K (Ed). Social media in social research: blogs on burning the boundaries. London, UK: NatCen Social Research.
Phu E. 2014. 8 insights into the market for ivory from social media in China. In: International Wildlife Trafficking Symposium Day 2 – A United for Wildlife event hosted by ZSL. https://www.youtube.com/watch?v=4cPkV0V3vzE [start time 6:09 minutes]. Viewed 14 Dec 2015.
Pijanowski BC, Villanueva-Rivera LJ, Dumyahn SL, *et al.* 2011. Soundscape ecology: the science of sound in the landscape. *BioScience* **61**: 203–16.
Poria S, Gelbukh A, Cambria E, *et al.* 2014. EmoSenticSpace: a novel framework for affective common-sense reasoning. *Knowl-Based Syst* **69**: 108–23.
Proulx R, Massicotte P, and Pepino M. 2014. Googling trends in conservation biology. *Conserv Biol* **28**: 44–51.
Schuetz J, Soykan CU, Distler T, *et al.* 2015. Searching for backyard birds in virtual worlds: internet queries mirror real species distributions. *Biodivers Conserv* **24**: 1147–54.
Thelwall M, Buckley K, Paltoglou G, *et al.* 2010. Sentiment strength detection in short informal text. *J Assoc Inf Sci Technol* **61**: 2544–58.
Wilde GR and Pope KL. 2013. Worldwide trends in fishing interest indicated by internet search volume. *Fisheries Manag Ecol* **20**: 211–22.
WWF (World Wildlife Fund). 2015. Safeguarding outstanding natural value: the role of institutional investors in safeguarding natural World Heritage sites from extractive activity. http://assets.wwf.org.uk/downloads/wwf_nwh_investor_report_a4_web_v2_1.pdf. Viewed 18 Jan 2016.
Żmihorski M, Dziarska-Pałac J, Sparks TH, *et al.* 2013. Ecological correlates of the popularity of birds and butterflies in internet information resources. *Oikos* **122**: 183–90.

Supporting Information

Additional, web-only material may be found in the online version of this article at http://onlinelibrary.wiley.com/doi/10.1002/fee.1260/suppinfo