

## TOOLS & METHODS

# Hiding in plain sight. A qualitative, text-based method for the analysis of websites

*Sarah Cummings, Bruce Boyes, Nyamwaya Munthali & Rocio Sanz*

### Abstract

This contribution develops a novel qualitative, text-based method for the analysis of websites. The texts of websites represent a huge resource for research that is rarely used in international development and is only very sparsely used in other fields. This method aims to make the most of the fact that websites provide a huge amount of 'authentic', unique, up-to-date, peer reviewed, topical data. Some of the challenges and opportunities of using the texts of websites for textual analysis in research are outlined, followed by a draft protocol for documenting individual pages which overcomes many of these challenges, making websites more accessible for qualitative research and, in particular, for critical discourse analysis. This method is based on a previous article which piloted a method to analyse the text of websites using critical discourse analysis (Cummings et al, 2025a). This method can also be employed more widely as a way to document websites which might be under threat of neglect and destruction.

Keywords: critical discourse analysis; discourse analysis; websites; internet; methodologies; policy analysis; tools

### 1 Introduction

*...we expect international organizations' websites to contain vital, carefully curated and up-to-date information about their goals, policies and activities, and we thus used websites to assess to what extent international organizations use the [Sustainable Development Goals] as a guiding framework and whether they are integrating policies. (Bogers et al., 2022: 136).*

*...[websites] are unique expressions of contemporary culture, and as such they constitute a huge repository of potential data about contemporary ways of doing and thinking of large groups of people across ethnic and national boundaries' (Pauwels, 2012: 247).*

There are many new tools emerging for the mining and analysis of website texts in an automated manner, partly for their use in artificial intelligence and machine learning. For example, Bogers and colleagues (2022) have undertaken a quantitative content analysis on the website texts of 159

international organizations to research the influence of the Sustainable Development Goals (SDGs) (UN, 2015) on policy. Websites are, however, not often used as sources for text for analysis using critical discourse analysis, despite the fact that they represent 'a huge repository of authentic data' (Mautner, 2005: 809) as well as being unique, up-to-date, current, self-reviewed, topical and linguistic resources (Fletcher, 2004). Although they have been used to analyse discourses of internationalization in higher education (Lewin-Jones, 2019), Sámi tourism websites (de Bernardi, 2019) and commercial trends for organic confectionary (Fernández-Vázquez, 2021) they are not widely used in international development, development studies and knowledge management. Therefore, the analysis of information on websites represent a novel approach and new resource in these fields. Thus, it represents data which is hiding in plain sight.

There is surprisingly little analysis of the texts on websites in qualitative research, using methodologies such as critical discourse analysis. Part of this might be because websites, apparently, invite scepticism and claims that online material is 'less real' than what may be learned from other data, such

as interviews or financial documents (Powell et al., 2016). Another possible explanation might be that there do not appear to be any protocols or methods for documenting individual webpages in the standard citation tools. This documentation is complex because individual pages often do not have clear authors, for example, and they are frequently updated. Without such a protocol, websites are like sand in the desert, changing almost every day and making them difficult to study in a consistent and systematic way. In view of this, we highlight how to document individual webpages and simultaneously outline some of the challenges and opportunities of using the texts of websites for textual analysis in research. This is followed by a draft protocol for documenting individual pages which, we consider, overcomes many of these weaknesses aforementioned, making websites more accessible for qualitative research. We argue that not all websites are appropriate for this research approach, and as such we present ongoing research and approaches of others to distinguish those which are.

To develop this tool for the textual analysis of websites, this article leans heavily on two earlier papers by the authors and other colleagues. First, a recent paper (Cummings et al., 2025a) analyses the websites of two communities of practice to identify their approaches to the decolonization of knowledge and epistemic justice that involved developing a draft protocol. It is the draft protocol developed there which we have further amended and are now making accessible to other researchers. Second, we are building on a similar type of 'Tools & Methods' paper in this journal which described the method of critical discourse analysis (Cummings et al., 2019). This method was originally developed for non-academics to support them in their use of critical discourse analysis but it has become a mainstream tool for textual analysis used by academics.

The method we are developing here is not only a tool for research. At a time when the knowledge on certain websites is apparently being destroyed for ideological reasons in a process going far further and faster than digital decay (Cummings et al., 2025b), the method also offers an approach to documenting and thus preserving the content of websites in a new international context where whole organizations and their websites are being put through the so-called woodchipper, such as what happened recently with multiple websites and publications of the US Agency for International Development (USAID) (Cummings et al., 2025b). This approach to systematic documentation thus supports knowledge retention in the widest sense, going beyond

current organizational practices as documented by Sanz and Howell (2021). Finally, as professionals who aim to bridge and research and practice, we think much of the content on websites is a huge resource hiding in plain sight, providing valuable insights into both research and practice, and that it is a waste of an enormous opportunity not to make better use of it in research.

## 2 Opportunities and challenges facing textual analysis of websites

According to recent analysis (Cardillo, 2025), there are currently 1,119,023,272 websites on the internet although only 17.3% are currently active. As of April 2024, some 67.1% of the global population or 5.44 billion people are users of the internet (Tambe & Jain, 2024). Globally, as of 2023, 71% of businesses have a website (Tambe & Jaun, 2024). As Kim argues, 'Every company sits on a goldmine of untapped potential—the data in its documents, including the letters, forms, checklists, compliance records and contracts it has with employees, customers, partners and suppliers' (Kim Technologies, 2024: unpaginated) and we argue that this is also the case for organizational and other websites. Pauwels argues that websites 'are unique expressions of contemporary culture, and as such they constitute a huge repository of potential data about contemporary ways of doing and thinking of large groups of people across ethnic and national boundaries' (2012: 247), also demonstrating the particular, potential contribution to international development which, more than any other domain, crosses ethnic and national boundaries. In a world where 'data is the new gold' (Shubladze, 2023, unpaginated), we need to look at new ways of using this data.

In the fields of international development and knowledge management for development (KM4D), international organizations, development agencies, research institutions, and international non-governmental organizations (NGOs) and community-based organizations all have websites. These websites feature detailed information about the organization, its staff and mission, in addition to blogs, educational content, policy advice, data resources, and platforms for engagement and discussion, depending on the purpose of the organization. In addition, communities of practice also have websites where members have discussions and place content. A previous paper by the authors and other colleagues (Cummings et al., 2025a) demonstrated that these websites can represent an authentic source of data and

information for many sorts of analysis, depending on an appropriate research methodology and research question. As was first noted by Mautner (2005) and was corroborated in the later article by Cummings and colleagues (2025a), websites give access to a huge amount of data and, thus, appear to be a valid source of data for critical discourse analysis. They also appeared to be an authentic source of data as described by Mautner (2005) because they found the same perspectives, multiple times. Given this repetition of perspectives, texts on website provide sufficient authenticity for critical discourse where even one document can provide evidence on a particular issue.

Websites of organizations have multiple roles: disseminating the innovative work and knowledge generated, providing the most accessible and comprehensive source of information on projects and findings, showcasing results and telling the stories of research journeys, and engaging with potential collaborators from around the world (Dillon-Shallard, undated). Against this background, a website is potentially the most comprehensive and representative source of data on the work, knowledge generated, projects and findings for development initiatives (Cummings et al., 20025a). Recent research indicates that digital content is being lost continually with more than one third of webpages that existed between 2013 and 2023 no longer being accessible, generally because an individual page was deleted or removed from a functional website (Chapekis et al., 2024). Chapekis and colleagues call this process 'digital decay' (2024: 2). According to the Digital Preservation Coalition (DPC) (2015), digital preservation, including the preservation of the content on websites, is 'a challenge of a generation' because of the fundamental vulnerability of digital resources due to their machine dependency. Although digital materials represent important resources for industry, commerce and trade, government, research, law, medicine, creative industries, cultural heritage and the media, their retention is under threat (DPC, 2015). According to the DPC, 'The greater the importance of digital materials, the greater the need for their preservation: digital preservation protects investment, captures potential and transmits opportunities to future generations and our own' (DPC, 2015: unpaginated). Websites are particularly vulnerable because of the rapid evolution of technology in which the software to create and access them can become obsolete, file formats change, and there is need for constant migration to newer storage mediums to prevent degradation.

Although website content can be ephemeral and suffers from the gradual degradation of digital decay, we argue that the standard introductory texts on websites are often relatively static. They have often been written and reviewed by multiple authors and, in this way, probably been 'formed, disseminated and legitimised within complex chains and networks of events' (Fairclough, 2013: 244-245), like policy documents. As Powell and colleagues note:

*The webpages of organizations are both a form of representation and a type of narrative. They entertain, persuade, express a point of view, and provide a means to organize collective action and economic exchange. Increasingly, webpages are the primary point of access between an organization and its environment (2016: 101).*

While introductory texts and foundational documents, such as memoranda of understanding, maybe static, new content is continually being added in terms of blogs and information about events. Announcements of events, such as seminars and conferences, will often be deleted automatically once they have taken place but blogs which appear over time as they are published will be kept online, often for many years and for as long as the website is in existence. Given the shifting nature of this content, a protocol for documenting it better is required, calling for the method that we are describing below. mple text inserted for illustration.

### **3 Websites worthy of serious, analytical attention**

We consider that not all websites deserve serious analytical attention. The University of New South Wales (UNSW) Foundation Studies Programme, for example, makes a distinction between websites which are suitable for academic use and those that are not. The university considers that the websites of government research organizations, university libraries and international research organisations, also identified as top-tier resources, are suitable for academic use (Table 1). It also identifies a range of mid-tier resources which should only be used with discretion and discernment, while it also argues that so called lower-tier sources, such as home-made websites and blogs, are unsuitable for academic use.

**Table 1.** University of New South Wales’ assessment of websites as academic sources.

Suitability for academic use	Top-tier sources	Mid-tier sources	Lower-tier sources
Types of websites	academic journal articles, text-books, government research, university libraries, international research organisations.	magazine articles, newspaper articles, academic blogs, local libraries, university sites.	home-made websites, personal blogs, wikis, social media sites, trade publications, shopping sites, essay sites.

Source: University of New South Wales

We would propose that only so called ‘top-tier websites’ with, potentially, bureaucratic sign-off procedures, would be suitable for serious analytical analysis. We note, however, that other, less formal websites may be a source of experiential knowledge and thus useful for other types of research. Against this background, we consider that websites of international organizations, bilateral development organizations, non-governmental organizations and others meet the criteria of top-tier websites. In Table 2, we have identified those websites which we consider are suitable for serious academic attention and have also listed potential questions and data sets, such as sub-websites, as inspiration. This list is not exhaustive and has been provided to give some examples.

**4 Method for the analysis and documentation of websites**

*And here, we must emphasize again the importance of the ‘first impressions’ for our study. Surprisingly, for such open-ended criteria, we achieved a high level of intercoder reliability. But it was not only mutual agreement. By matching our impressions, we could generate important and rich data to our research, such as how ethnic groups perceive and construct their digital presence. (Caneiro & Johnson, 2014: 14).*

To develop this method, we built on the work of others who have analysed websites. Carneiro and Johnson (2014), for example, analyse the images on websites, using the first four steps of Pauwels’ (2012) six step method called a ‘multimodal framework for analyzing websites.’ This framework is designed for research on cultural differences between countries or ethnicities but is apparently also suited to track differences between organizational cultures. These steps of the framework comprise:

- Step 1: The qualitative preservation of the first impression.
- Step 2: The quantitative inventory of salient and absent features.
- Step 3: In-depth analysis of content and formal choices (written, visual, and sonic signifiers, layout).
- Step 4: Stand points of view and implied voice and audience.
- Step 5: Information organization and navigation.
- Step 6: Contextual analysis, such as affordances and constraints of the technological platform.

Although the Pauwels’ (2012) method with all six steps has previously been used for in-depth analysis of a smaller number of websites, Carneiro and Johnson (2014) only use the first four steps because they were analysing multiple websites. They consider that the fifth step (information organization) is very time consuming because most websites are ‘very complex, sophisticated, and with many layers and possibilities for navigation’ (2014: 6) and that this step is not crucial for answering their research questions, namely how ethnic museum websites in the USA visually construct ethnic identity and represent cultural heritage. In fact, they consider that the fifth step is more suitable for the analysis of technical qualities. We argue, however, that the fifth step is of vital importance given the need to document websites accurately in a situation where they are not only suffering gradual degradation but are actually under threat for ideological reasons so we include this step in our three phased method. Based on the work of Pauwels (2012) and Caneiro and Johnson (2014), we use our experiences of analysing communities of practice websites (Cummings et al, 2025a) and of critical discourse analysis (Cummings et al., 2019; Seferiadis et al., 2021), based on the work of Fairclough (2012a, 2012b) to develop a methodology which reflects previous experience and is tailored the contemporary threats websites and digital information face. In addition, we combine this with the method of critical discourse analysis in three phases. We illustrate this method with the example of

the paper on communities of practice (Cummings et al, 2025a).

*Phase I: Selection of research topic and providing an overview of past discourses*

Step 1: Select a research topic that relates to a social question that can be productively approached with a focus on text on a single or multiple websites.

Example: We wanted to study how decolonisation of knowledge was happening in academia and what this looks like in practice.

Step 2: Which discourses have already been identified in the area you are investigating?

Example: In this case, we had a framework of epistemic justice (Cummings et al, 2023) which we wanted to test for its adequacy in practical use but also to see how we could improve it.

*Phase II: Selection and analysis of texts on the website*

Step 1: Select websites appropriate to the object of research

Example: Two communities of practice were identified that included substantial information on their websites, publicly available. This made them a fairly unique resource as many communities of practice’s deliberations are only available to members.

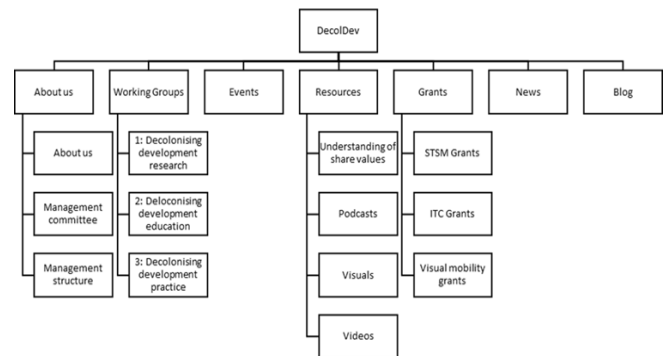
Step 2: The qualitative preservation of the first impressions  
This step has two parts, derived from Pauwel’s step one, namely, first, categorizing the ‘look and feel’ of the website at a glance and second afford your affective to the website: how does it make you feel: are you attracted, intrigued or puzzled by it? As Pauwels explains:

*These first reflections need to be recorded while they remain spontaneous and they are important to feed a reflexive attitude, which implies the conscious reception of a website as a ‘meeting of cultures’ between producers, intended audiences and researchers. Such reflexivity is also required throughout the research process to help understand the reactions of other people, who have not studied the website as a focus of research. (2012: 253).*

Although we have no specific example of this approach, we think it could be a valuable starting place for analysis. In this process, we are proposing that this is a preliminary exercise.

Step 3: Developing an overview of the page structure of the websites.

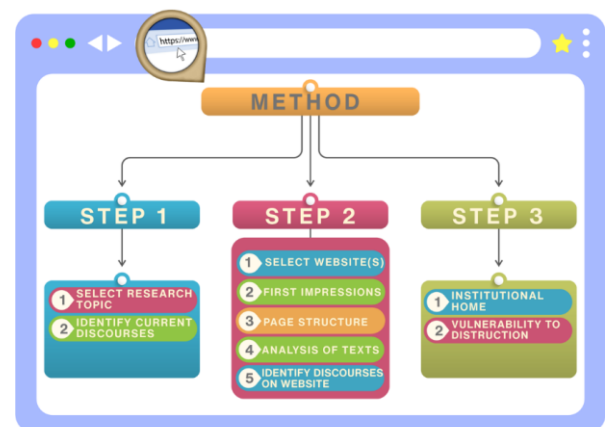
**Figure 1.** An overview of the structure of pages on a website



Source: Cummings et al, 2025.

Example: We mapped the page structure of the COST Action DecolDev website (Figure 1) which has 7 main pages and 13 subpages as a crucial step in documenting the website structure. This step is consistent with part of Pauwels’ Step 5 which involves ‘Analysis of Information Organization’ but, here, the purpose is different, namely our purpose to include more rigour in the textual analysis by making it clearer where the text is found and also capturing images of the text, including a date stamp. Therefore, this stage serves the purpose of documenting the website, its pages and text in a situation where the websites might no longer be in existence, either because an initiative has finished or the website has been taken down for any other reason.

**Figure 2.** Overview of the method in phases and steps.



Source: Authors

**Table 2.** Types of organizations and possible use of data to be found on their websites

Types of organizations	Illustrative examples	Some illustrative examples of what could be analysed
International organizations	UN Food and Agriculture Organization Connect website: <a href="https://www.fao.org/connect-private-sector/en/">https://www.fao.org/connect-private-sector/en/</a>	UN discourses in the private sector for the agricultural sector
	UN International Fund for Agricultural Development <a href="https://www.ifad.org/en/indigenous-peoples">https://www.ifad.org/en/indigenous-peoples</a>	UN discourses on Indigenous peoples and their knowledge
Bilateral organizations	UN Children’s Fund <a href="https://knowledge.unicef.org/">https://knowledge.unicef.org/</a> <a href="https://www.government.nl/topics/development">https://www.government.nl/topics/development</a>	UN discussions relayed to child nutrition Discourses of Dutch development aid
	OXFAM International Development Enterprises Water Aid	Approaches to development and related trending terms
Research institutes	Institute of Development Studies, University of Sussex, UK	Trends in Development studies research and key discourses
	Bangladesh Institute of Development Studies (BIDS)	Trends in Development studies research and key discourses
Learned societies	Sub-website on <a href="https://www.eadi.org/development-studies/decolonising-knowledge-for-development">https://www.eadi.org/development-studies/decolonising-knowledge-for-development</a>	Approaches to decolonisation of knowledge
Communities of practice	Knowledge Management for Development (KM4Dev) <a href="https://dgroups.io/g/km4dev">https://dgroups.io/g/km4dev</a>	Discourses around knowledge for development
	Health Information for All (HIFA) <a href="http://www.hifa.org">www.hifa.org</a>	Discourses related to global and public health
	Dgroups Foundation <a href="http://www.dgroups.info">www.dgroups.info</a>	Discourses related to climate change

Source: Authors

**Step 4: Analysis of the texts**

This is the main step in the research process, namely analysis of texts manually or using qualitative analysis software at the level of: individual words and phrases; how the words and phrases relate to each other in the text; and the priority given to different themes.

Example: In this step, we analysed the texts of the websites in detail to look for themes related to epistemic justice and decolonization of knowledge which was the framework we had identified in Phase 1, Step 2.

We needed to develop a protocol for citing the content from the different webpages. We did this by naming the website, identifying the page, and the date on which the page was consulted, aiming to make the citations as specific as possible. For example, if we consulted the ‘About Us’ page on the DecolDEV website on 17 July 2024, we referred to it as follows (DecolDEV: About Us, 17 July 2024). Given the overview of the page structure and the date, we felt that this

was specific enough without giving the URL every time because the main URL remains the same. These references are not to be found in the list of references. For all other documents that we consulted on the websites, such as blogs and memoranda of understanding, we used traditional referencing protocols and these documents can be found in the list of references. Given that there is no accepted protocol for identifying them in the reference list, we marked them with an asterisk (\*). This, for example, generated the references:

\*Dey, S. (2023). Ecocentric Pedagogies and Green Scholarships: Towards Green Academia. *Convivial Thinking*, February 11, 2023.

<https://convivialthinking.org/index.php/2023/02/11/ecocentric-pedagogies/#more-2097>.

\*Decolonising Development. (2021). Statement of the Threat of Academic Authoritarianism. April 12, 2021. <https://decolonise.eu/statement-on-the-threat-of-academic-authoritarianism/>.

**Step 5: Identify discourses in the text, based on the prior discourses identified in Phase 1.**

Example: We used the discourses around epistemic justice and decolonization to code the content of the websites, aiming to find areas that converged and diverged from the framework.

*Phase 3: Describe how the website was created and how this relates to the discourses identified*

Step 1: Which institution developed the website and does the activity appear to be time-bound?

Example: These websites were created specifically to present and showcase the findings of the two communities of practice to members and also to the outside world. The COST website was created for a directly time-bound activity, namely the COST Network while Convivial Thinking is not time-bound but appears to be inactive (September 2025) while the COST Action ended in September 2024.

Step 2: Is the website vulnerable to being lost?

Example: It appears that Convivial Thinking website may be the most vulnerable because the administrators will need to keep up payments for a website domain that they no longer appear to be using actively, while the COST Action website is supported by the European Commission which hosts many future and past COST Actions and may be less vulnerable to loss. To illustrate this with an example, the Knowledge Management for Development (KM4Dev) community's knowledge sharing toolkit website was lost because it was not clear which individual had the responsibility for renewing it. In another example, access to KM4Dev's main website was almost lost when the domain was being paid for automatically by the hosting organization's information technology department, unknown to other members of the organization in another department. These two examples also demonstrate the vulnerability of websites, even within organizations. Figure 2 provides an overview of the method in terms of phases and steps.

## 5 Conclusions

With this article, we present a method to undertake textual analysis of websites in the field of international development studies and knowledge management for development. We hope that the tool can be further improved by use so we would be grateful for feedback from anyone who uses it. We hope the method will be useful for academics, researchers and practitioners who are trying to understand discourses through the analysis of discourses.

## Acknowledgements

We acknowledge with thanks review of a previous draft by Leah de Haan. Remaining errors are our own.

## About the authors

*Sarah Cummings* is a Research Fellow at Wageningen University, the Netherlands. She is also the Editor-in-Chief of the *Knowledge Management for Development Journal*. She was previously Co-Director for the IKM Emergent Research Programme and is chair of the KM4Dev community. She holds a PhD from the VU University Amsterdam, a Masters' from Radboud University Nijmegen, and a Bachelor's from SOAS. Email: [sarahcummingswork@gmail.com](mailto:sarahcummingswork@gmail.com).

*Bruce Boyes* is editor, lead writer, and a director of the award-winning *RealKM Magazine*. He holds a Master of Environmental Management with Distinction, and his work in cross-domain knowledge integration and knowledge co-creation for sustainable development is highly regarded. His expertise and experience include knowledge management (KM), environmental management, project management, stakeholder engagement, teaching and training, communications, research, and writing and editing. Through RealKM Magazine, Bruce has been very active in advancing the decolonization of knowledge and KM. An impetus for this has been eight years of living and working in China, where he became acutely aware of the extent to which the global knowledge base is biased against the global South, including learning about Chinese KM approaches that offer advantages over dominant Euro-American approaches. Email: [bruce@bruceboyes.info](mailto:bruce@bruceboyes.info)

*Nyamwaya Munthali* holds a PhD in Development Studies from Wageningen University and a PgD in Teaching Methodology from the University of Lusaka. She lectures in Communication and Development at the University of Zambia and previously served at Cavendish University as Acting Head of Department and member of the PhD and Publications Board. Her research explores digital communication, technological innovation, and the decolonisation of knowledge. She sits on the editorial board of the *Knowledge Management for Development Journal* and has over eight years' experience in Zambia's agriculture sector, focusing on mobile innovation, livelihoods, enterprise development, and project management. Email: [nyamwaya.munthali@unza.zm](mailto:nyamwaya.munthali@unza.zm)

Rocio Sanz is a Senior Knowledge Broker and Consultant with a robust background in communication, learning, and knowledge management. For over 20 years, she has worked with UN agencies—including UNICEF, IOM, IFAD or UNDP—as well as non-governmental organizations and governments, serving, mentoring, and coaching organizations, teams, and individuals. She specializes in strengthening knowledge strategies, intelligence, and frameworks for development, with a particular focus on knowledge transfer and retention. Rocio excels in capacity building and in designing customized, innovative tools for effective knowledge sharing and generation. Her expertise extends to working with communities and networks, fostering collaboration, learning, and the exchange of tacit knowledge. She holds a Master of Science in Information Architecture and Knowledge Management from Kent State University (USA), a Master of Arts in Communications and International Relations from Complutense University (Spain), and she is currently a PhD candidate focusing on knowledge retention by the Knowledge, Technology and Innovation (KTI) of Wageningen University and Research (WUR) in the Netherlands. As a core member of KM4Dev, she also serves as a Senior Editor for the *Knowledge Management for Development Journal*.

Email: sanz\_rocio@hotmail.com

## References

- de Bernardi, C. (2019). Authenticity as a compromise: a critical discourse analysis of Sámi tourism websites. *Journal of Heritage Tourism*, 14(3), 249–262. <https://doi.org/10.1080/1743873X.2018.1527844>
- Bogers, M., Biermann, F., Kalfagianni, A., & Kim, R. E. (2022). Sustainable Development Goals fail to advance policy integration: A large-n text analysis of 159 international organizations. *Environmental Science and Policy*, 138, 134–145. <https://doi.org/10.1016/j.envsci.2022.10.002>
- Cardillo, A. (2025) How Many Websites Are On The Internet? (2025). *Exploding Topics*, 28 May 2025. <https://explodingtopics.com/blog/how-many-websites-on-the-internet>
- Carneiro, L., & Johnson, M., (2014). Quantitative and qualitative visual content analysis in the study of websites. In *Sage Research Methods Cases Part 1*. SAGE Publications, Ltd., <https://doi.org/10.4135/978144627305013517800>
- Chapekis, A., Bestvater, S., Remy, E., & Rivero, G. (2024). When Online Content Disappears: 38% of webpages that existed in 2013 are no longer accessible a decade later. *Pew Research Center Report*.
- Cummings, S.J.R., Munthali, N., & T. Sittoni (2025a) Epistemic justice as a ‘new normal’? Interrogating the contributions of communities of practice to decolonization of knowledge. *Sustainable Development* 33(3): 3228-3245.
- Cummings, S.J.R., White, N. & Boyes, B. (2025b) USAID and the new burning of the books in digital and ideological epistemicide. A call to action. *RealKM Magazine* 27 February. <https://realkm.com/2025/02/27/usaid-and-the-new-burning-of-the-books-in-digital-and-ideological-epistemicide-a-call-to-action/>.
- Cummings, S., De Haan, L., & Seferiadis, A. A. (2020). How to use critical discourse analysis for policy analysis: a guideline for policymakers and other professionals. *Knowledge Management for Development Journal*, 15(1), 99-108.
- Digital Preservation Coalition (2015). *Digital Preservation Handbook*, Second Edition.
- Dillon-Shallard, D. (undated). Virtual presence, global impact: the indispensable benefits of a website for research institutes. *Butterfly blog*. <https://butterfly.com.au/blog/website-for-research-institutes/>
- Fairclough, N. (2012a) The dialectics of discourse. Unpublished paper. <https://www.sfu.ca/cmns/courses/2012/801/1-readings/Fairclough%20Dialectics%20of%20Discourse%20Analysis.pdf> (Accessed 11 May 2016)
- Fairclough, N. (2012b) Critical Discourse Analysis. In: *The Routledge Handbook of Discourse Analysis*, edited by James Paul Gee and Michael Handford, 9-21. Abingdon and New York: Routledge.
- Fairclough, N. (2013). *Critical discourse analysis: The critical study of language*. Abingdon: Routledge.
- Fernández-Vázquez, J. S. (2021). Selling organic candy: multimodal critical discourse analysis of commercial websites. *British Food Journal*, 123(10), 3277-3292.
- Fletcher, W. H. (2004). Making the Web More Useful as a Source for Linguistic Corpora. *Language and Computers*, 52, 191–206.
- Kim Technologies. (2024). Is your data hiding in plain sight? <https://blog.kimdocument.com/blog/is-your-data-hiding-in-plain-sight>
- Mautner, G. (2005). Time to get wired: Using web-based corpora in critical discourse analysis. *Discourse & Society*, 16(6), 809-828.
- Pauwels, L. (2012). A multimodal framework for analyzing websites as cultural expressions. *Journal of Computer-Mediated Communication*, 17, 247–265. doi: <http://dx.doi.org/10.1111/j.1083-6101.2012.01572>
- Powell, W.W, Horvath, A., & Brandtner, C. (2016) Click and mortar: Organizations on the web. *Research in Organizational Behavior* 36: 101-120. <https://doi.org/10.1016/j.riob.2016.07.001>.
- Sanz, R. & Hovell, J. (2021). Knowledge retention framework and maturity model: improving an organization or team’s capability to retain unique and critical

- knowledge. *Knowledge Management for Development Journal* 16(1): 8-27.
- Seferiadis, A. A., de Haan, L., & Cummings, S. (2021). Feminist Critical Discourse Analysis of Ecopreneurship as an Instrument for Sustainable Development Grand Narratives and Local Stories. In *Environmental Sustainability and Development in Organizations* (pp. 1-17). CRC Press.
- Shubladze, S. (2023) How To Make Use Of The New Gold: Data. *Forbes*, 27 March 2023.  
<https://www.forbes.com/councils/forbestechcouncil/2023/03/27/how-to-make-use-of-the-new-gold-data/>
- Tambe, N., & A. Jain (2024) Top website statistics for 2024. *Forbes*, 28 June 2024.  
<https://www.forbes.com/advisor/in/business/software/website-statistics/?via=hemin74>
- UN. (2015). *Transforming our world: the 2030 Agenda for Sustainable Development*.  
<https://sustainabledevelopment.un.org/post2015/transformingourworld/publication> (Accessed 15 January 2019)