

Unpacking IPCC and IPBES Reports

Shweata N. Hegde[‡], Ayush Garg[§], Peter Murray-Rust[†], Daniel Mietchen^{¶, #, «}

[‡] Regional Institute of Education, Mysuru, India

[§] Global Indian International School, Singapore, Singapore

[†] University of Cambridge, Cambridge, United Kingdom

[¶] Ronin Institute for Independent Scholarship, Montclair, United States of America

[#] Institute for Globally Distributed Open Research and Education (IGDORE), Jena, Germany

[«] Leibniz Institute of Freshwater Ecology and Inland Fisheries, Berlin, Germany

« Freie Universität Berlin, Berlin, Germany

Corresponding author: Daniel Mietchen (daniel.mietchen@ronininstitute.org)

Abstract

Humanity is facing a set of existential challenges, including the handling of the parallel and interconnected crises of climate change and biodiversity loss. In an effort to address these challenges, international bodies like the Intergovernmental Panel on Climate Change ([IPCC](#)) and the Intergovernmental Science-Policy Platform on Biodiversity and Ecosystem Services ([IPBES](#)) have been created. These bodies are producing a series of reports that compile scientific information on specific aspects of climate and biodiversity research and discuss policy options and other social implications.

So far, these reports have been provided in formats that make it hard to mobilize the knowledge encapsulated in them. Our contribution demonstrates technical workflows for achieving such mobilization by mining the reports. Development of these workflows is spearheaded by young researchers from the SemanticClimate team of interns based at the National Institute of Plant Genome Research (NIPGR) in New Delhi, India, and volunteers from all over the world. The tool chain includes methods for cleaning up the formatting, for extracting and processing raw text, tables and figures and annotating them semantically with the help of controlled vocabularies, ontologies and Wikidata. The semantic information and the mined information can then be combined in a way that iteratively improves both, eventually resulting in versions of the reports wherein entities like species, countries or references are semantically marked up and rendered in responsive formats.

Our framework supports multilingual and specialist interests (e.g., endangered species, plant chemistry), and we will also briefly discuss how the use of standard open licensing could further contribute to mobilizing information from the reports. We are keen to work with other groups sharing these interests as well as with the teams involved in producing the reports.

Keywords

climate change, biodiversity, semantic publishing, Wikidata

Presenting author

Shweata N. Hegde

Presented at

TDWG 2022

Acknowledgements

This work was supported by the Alfred P. Sloan Foundation under grant number G-2021-17106.

Conflicts of interest