

# Notes on the CBSpeeches Dataset

Emanuele Campiglio<sup>a,b,c</sup>, Jérôme Deyris<sup>d</sup>, Davide Romelli<sup>e</sup>, and Ginevra Scalisi<sup>f</sup>

<sup>a</sup>*University of Bologna, Department of Economics, Italy*

<sup>b</sup>*CMCC Foundation - Euro-Mediterranean Center on Climate Change, Italy*

<sup>c</sup>*LSE Grantham Research Institute on Climate Change and the Environment, United Kingdom*

<sup>d</sup>*Sciences Po Paris, Centre d'Études Européennes, France*

<sup>e</sup>*Trinity College Dublin, Ireland*

<sup>f</sup>*Scuola Superiore Sant'Anna, Italy*

This version: January 2025

## Abstract

The CBSpeeches Dataset is the dataset of central banker speeches collected for Campiglio et al. (2025). To the best of our knowledge, this dataset represents the most comprehensive collection of public speeches by central bankers. It was created by complementing the [Central bankers' speeches dataset](#) compiled by the Bank of International Settlement (18,045 speeches) with a systematic scraping of central banks' websites (15,435 new speeches) and archive work (2,007 additional speeches).

The dataset features 35,487 unique speeches from 131 central banks, for the period going from the beginning of January 1986 (date of the first online speech by a central bank) to the end of December 2023. Among those speeches, 5,347 were provided in a non-English language and were translated using Microsoft Translation services.

The dataset is available at [cbspeeches.com](https://cbspeeches.com) and will be periodically updated by the author.

When you use the index and its associated data, please acknowledge the following papers as the data sources:

- Campiglio, E., Deyris, J., Romelli, D., Scalisi, G., 2025. [Warning words in a warming world: Central bank communication and climate change](#). *Grantham Research Institute on Climate Change and the Environment*, Working Paper No. .

# 1 Codebook

- ✓ **URL:** URL from which the speech has been originally extracted, if available.
- ✓ **PDF:** Pdf link from which the speech has been originally extracted, if available.
- ✓ **Title:** Speech title.
- ✓ **Subtitle:** Speech subtitle, if available.
- ✓ **Date:** Date of the speech.
- ✓ **Authorname:** Speaker name.
- ✓ **Role:** Official role of the speaker (Governor, deputy, etc) on the day the speech was delivered.
- ✓ **Gender:** Gender of the speaker.
- ✓ **CentralBank:** Name of the central bank/institution delivering the speech.
- ✓ **Country:** Country associated with the central bank.
- ✓ **text\_original:** Text of the speech. This is the unprocessed text that we have extracted from the speeches.
- ✓ **Filename:** Name of the file in our database.
- ✓ **Language:** The official language in which the speech was delivered/downloaded.
- ✓ **Source:** Information on the source of the speeches (archives, central bank website, BIS dataset).