Addressing Data Gaps in Sustainability Reporting:

A Benchmark Dataset for Greenhouse Gas Emission Extraction

TL;DR: We provide gold-standard CO2 emission data for company reports!

1. Background & Problem:

Company-level greenhouse gas (GHG) emissions are

essential but existing datasets are often:

- \rightarrow Fragmented
- \rightarrow Inconsistent
- \rightarrow Lack transparent methodology

nvironmental managemen	Table ENV-3				
nis section contains data related to the	environmento	ıl performar	nce of Allian	z Group.	Energy consumption
ıble ENV-1					As of December 31
MS Coverage					Energy consumption fro office buildings
s of December 31		2022	2021	2020	Energy consumption fro
hare of employees in scope of our nvironmental management system (EMS)	%	95	96	97	data centers Total energy consumpt
ıble ENV-2					Energy consumption fro buildings per employee
reenhouse gas emissions					Energy reduction in offi
s of December 31		2022	2021	2020	Target -20% by year-end 2025
cope 1 – Direct GHG emissions	t CO ₂ e	30,953	28,699	28,714	
cope 2 – Indirect GHG emissions market based)	t CO ₂ e	30,490	54,689	100,722	Table ENV-4 Energy sources
cope 2 – Indirect GHG emissions ocation based)	t CO ₂ e	138,339	599	180,826	As of December 31
cope 3 – Other indirect GHG emissions	t CO ₂ e	92,467	55,359	73,916	Electricity
cope 1–3, GHG emissions total	t CO ₂ e	153,910	138,746	203,352	Fossil fuels
otal GHG Emissions per employee Overall GHG reduction per employee ince 2019 pard target -50% by end-year 2025	t CO2e/empl %	<u>1.0</u> -57	0.9 -60	<u>1.4</u> -42	Long distance heating Other sources (incl. ene own sources including p internal waste heat)

t CO₂e







2. Our approach: Complementing large but erroneous datasets with small gold-standard validation set:

Carbon Credits

 \rightarrow Define objective and clear extraction criteria

 \rightarrow Employ hybrid LLM-human extraction pipeline

 \rightarrow Evaluate extractions with two-staged expert adjudication and discussion process

3. Data collection

Decision rules: Emission values are

considered in our data if they are reported

- 1. for the **whole** ("consolidated") company
- 2. in **absolute** tonnes of CO_2 or CO_2e
- as a **total**, not subcategory 3.



4. according to the operational boundaries of

the **scopes** (according to GHGP)

controlled by the reporting organization" e.g., company facilities

indirect GHG emissions that are related to a company's purchases of electricity, steam, heating, or cooling from utility providers

chain emissions including the GHG releases of purchased goods, business travel, use of sold products and investments



Preliminary results:

Scope	Average reportings per report
1	3.33
2lb	2.94
2mb	1.23
3	1.71

4. Next Steps

What we do:

 \rightarrow LLM/pipeline error analysis

 \rightarrow Develop automated extraction pipeline

 \rightarrow Extend extraction pipeline to additional metrics

(e.g. greenwashing/transparency indicators)

What you can do with our data:

 \rightarrow Benchmarking: Compare other models, prompts

 \rightarrow Information extraction experiments

(automated/human/hybrid)

 \rightarrow Substantive analysis

Jacob Beck1,3, Anna Steinberg1,3, Andreas Dimmelmeier1, Laia Domenech Burin1, Emily Kormanyos2, Maurice Fehr2, and Malte Schierholz1,3 1LMU Munich, Department of Statistics, 2Deutsche Bundesbank, Research Data and Service Centre 3Munich Center for Machine Learning (Paper under review) Contact: jacob.beck@lmu.de + anna.steinberg@lmu.de

All views expressed in this poster are personal views of the authors and do not necessarily reflect the views of Deutsche Bundesbank or the Eurosystem.

