

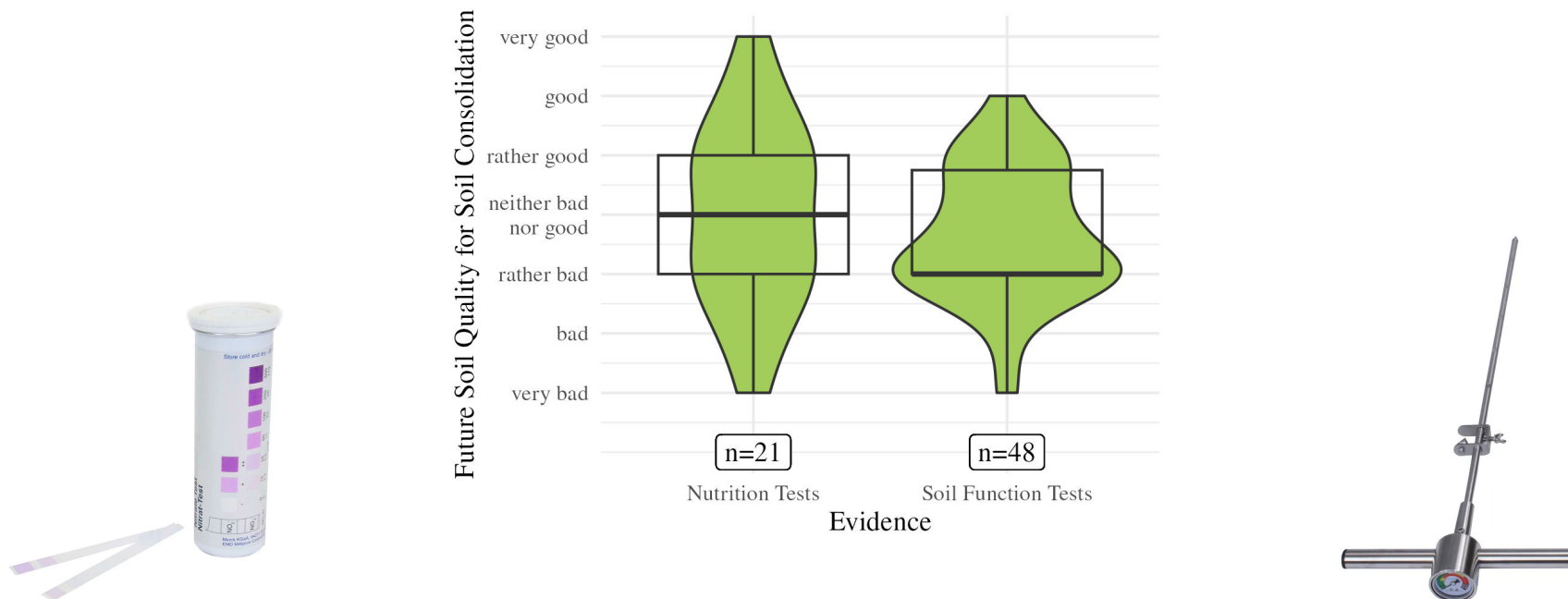
Closing the Granular Data Gap

A cross-domain perspective on information architectures in the context of climate change and nature depletion

Dr. Felicitas Sommer (LMU & TU Munich)

Two realities: agricultural nutrition vs. soil function measurement

Survey Question: what is the soil health status in 10 years?



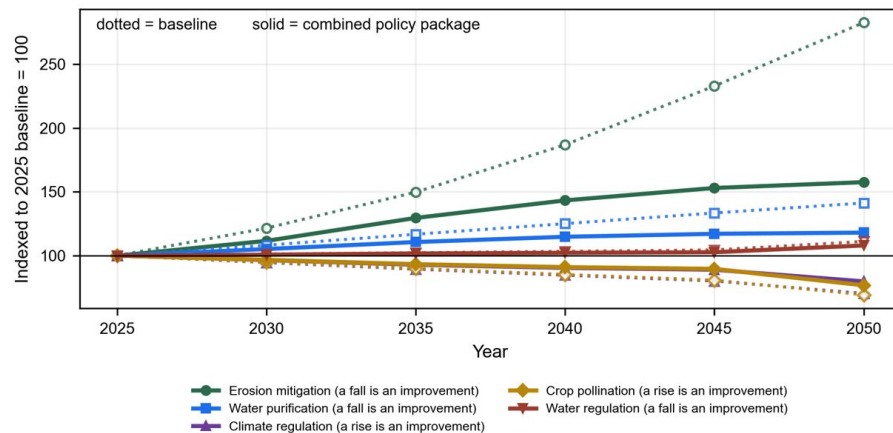
How land-use policies impact GDP tackling erosion mitigation

Based on OPEN IEEM+ESM Platform: Integrated Environmental-Economic Modeling (IEEM)

“Nature-Based Solutions from Fiscal Liability to National Asset: Internalizing Ecosystem Services in Malaysia”

A recursive-dynamic Computable General Equilibrium model of Malaysia is coupled to spatially explicit land use and land cover change modeling and to eight Ecosystem Services (ES).

It calculates GDP loss in 2060 for different scenarios of land use and land cover change.



Source: <https://virtual.oxfordabstracts.com/event/75057/submission/229>

Onil Banerjee1, Martin Cicowiez2, Mani S. Muthukumara3, Hector Pollitt4, Pralhad Burli1, Kenneth J. Bagstad5, Renato Vargas6, Carl Obst7, Disussant: Sebastian Roux8
1 RMGEO Consultants Inc., Fredericton, Canada. 2 National University of La Plata, La Plata, Argentina. 3 World Bank, Hanoi, Vietnam. 4 University of Cambridge, Cambridge, United Kingdom. 5 Data Foundation, Washington DC, USA. 6 CHW Research, Guatemala City, Guatemala. 7 IDEEA Group, Melbourne, Australia. 8 French National Institute of Statistics and Economics Studies (INSEE), Montrouge, France

Spatial governance data

Why it matters

1. Land-use regulation and the climate adaptation gap

Conceptual background

2. The systemic climate adaptation framework

Method

3. Spatial information from corporate sustainability reports
4. Regulatory information from local binding land-use plans

Conclusion and Outlook

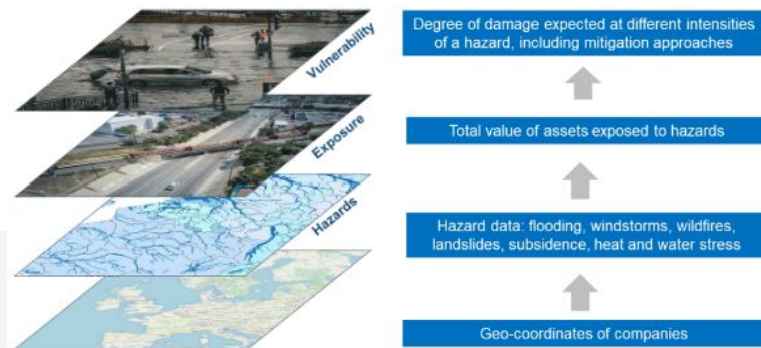
5. Land management information systems and federated data architectures

1. Land-use, climate adaptation, and risk assessment

The relevance of adaptation measures for climate risk assessment

- **Adaptation measures** are important for climate risk assessments (European Central Bank, 2024)
- More accurate protection data is necessary (Rohmer et al. 2026)
- For reliable datasets on coastal flood protection measures a transdisciplinary community-based approach is needed (de Plaen 2025)
- **Adaptation gap:** implementation of adaptation measures is hindered by constraints including institutional capacity, social inertia, and cost (Aerts et. al. 2024)

Stylised composition of physical risk indicators



Source: ECB, Climate change-related indicators, Methodological report, January 2023, based on the United Nations Office for Disaster Risk Reduction (UNDRR) terminology.

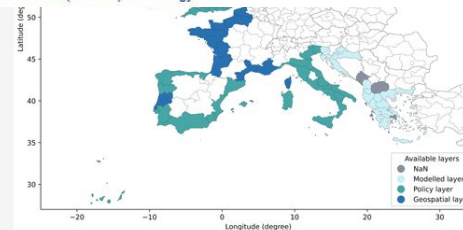


Figure 1 Data availability overview of COASTPROS-EU, representing the best-available coastal protection standards in Europe per NUTS2 region for three typologies of layers: geolocated coastal defences, policy standards, and modelled defence standards.

Data Gap: Institutional response and preparedness to climate change

Knowledge gap:

How well are adaptation measures implemented and enforced?

What do we **not** know about the

Adaptation gap?

nature water

Perspective

<https://doi.org/10.1038/s44221-024-00274-x>

Exploring the limits and gaps of flood adaptation

Received: 4 December 2023

Accepted: 10 June 2024

Published online: 16 July 2024

 Check for updates

Jeroen C. J. H. Aerts^{1,2}, Paul D. Bates³, W. J. Wouter Botzen⁴,
Jens de Bruijn^{1,4}, Jim W. Hall⁵, Bart van den Hurk^{1,2}, Heidi Kriebich⁶,
Bruno Merz^{1,2}, Sanne Muis^{1,2}, Jaroslav Myslak⁷, Eric Tate⁸ &
Frans Berkhout⁹

Flood adaptation measures such as levees, flood-proofing structures, nature-based solutions and flood insurance are essential to cope with the growing flood risk caused by climate change and urban development into flood-prone areas. However, many communities in flood zones are ill-protected because the implementation of adaptation measures is hindered by a variety of constraints to adaptation, including the cost, limitations on institutional capacity and societal inertia. When adaptation efforts fall short relative to a desired level due to a combination of constraints, it results in an 'adaptation gap'. Here we present a risk-based framework to systematically examine the technical, social and behavioural constraints that contribute to adaptation gaps. We argue that, without overcoming these constraints, adaptation gaps will widen under climate change, exposing increasing populations to heightened flood risk. This may then require more radical actions including relocation, as risks become intolerable. We argue that quantitative flood risk assessments must consider constraints and adaptation gaps systematically, especially where they may lead to flood adaptation limits. Without assessing these dynamic relationships, flood managers may overestimate the efficacy of flood adaptation measures and underestimate the unequal distribution of flood risks.

Adapting to floods is crucial to lower the growing threat posed by flood-related hazards¹. Throughout time, societies have utilized diverse strategies to cope with flooding, including seasonal flood recession agriculture, positioning settlements on locally elevated ground and protection with dykes and levees. More recently, quantified flood risk analysis, early warning systems and flood insurance have provided further strategies to mitigate risks linked to river, pluvial and coastal flooding. Yet, numerous constraints stand in the way of flood adaptation efforts². Factors like high costs of protection or poor governance limit the adoption of these measures, leading to increased risk³. Climate change will further increase the frequency and intensity of flooding, and

falling to intensify flood adaptation efforts will thus lead to additional risks and increased demands for financial compensation⁴. Although compensation schemes and insurance have a clear role to play in flood management, they are only partial responses to these trends⁵. Since current and future adaptation efforts are bound by constraints, in many cases we can observe an 'adaptation gap'. This gap is the risk that persists after subtracting (1) the risk associated with a societally desired level of adaptation from (2) the risk associated with actual, constrained adaptation measures that are taken⁶. Some argue that an economically optimal desired level of adaptation can be arrived at via cost-benefit analysis (CBA) and the rational allocation of available

¹Institute for Environmental Studies, VU University Amsterdam, Amsterdam, The Netherlands. ²Deltares, Delft, The Netherlands. ³University of Bristol, Bristol, UK. ⁴International Institute for Applied Systems Analysis, Laxenburg, Austria. ⁵University of Oxford, Oxford, UK. ⁶German Research Centre for Geosciences GFZ, Potsdam, Germany. ⁷University Potsdam, Potsdam, Germany. ⁸Euro-Mediterranean Center for Climate Change, Venice, Italy. ⁹Princeton University, Princeton, CA, USA. *King's College London, London, UK. ✉ e-mail: jeroen.aerts@vu.nl

2. The systemic climate adaptation framework (SEI)

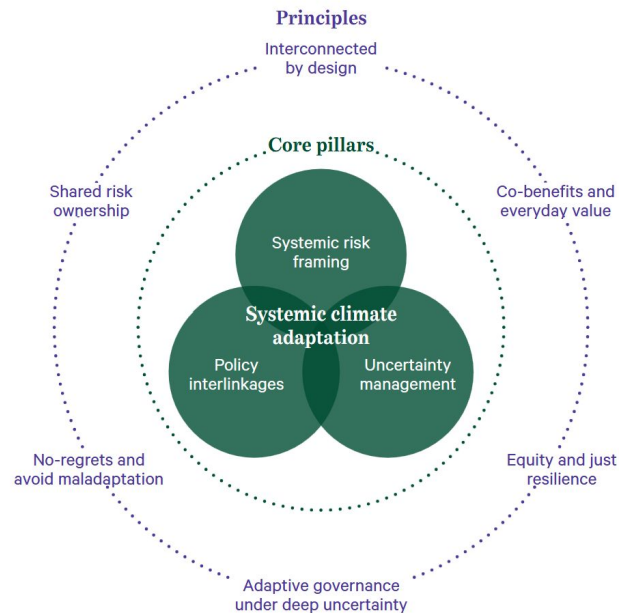
Stockholm Environment Institute 2026

The Systemic Climate Adaptation Framework

Systemic adaptation requires governance arrangements that reflect the way risks propagate in reality, rather than the way bureaucratic boundaries are drawn.

Responsibilities for assessment, decision-making and implementation are often fragmented, leading to gaps, overlaps and **“orphaned”** risks.” (SEI, P.6)

Figure 1. The framework consists of three conceptual pillars that support six principles for guiding policy and governance.



Source: Mikaelsson, M. and Shu, M. (Stockholm Environment Institute).

The principles of the systemic climate adaptation framework

Principle 1	interconnected by design
Principle 2	shared risk ownership
Principle 3	co-benefits and everyday value
Principle 4	no-regrets and anti-maladaptation
Principle 5	equity and just resilience
Principle 6	adaptive governance under deep uncertainty

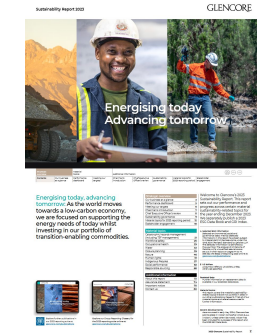
The principles of the systemic climate adaptation framework

Principle 1	interconnected by design
Principle 2	shared risk ownership
Principle 3	co-benefits and everyday value
Principle 4	no-regrets and anti-maladaptation
Principle 5	equity and just resilience
Principle 6	adaptive governance under deep uncertainty

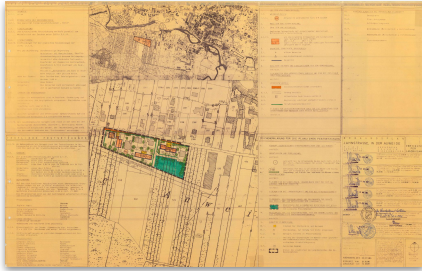
Cross-sectoral gap: private sector governance and land-use planning

High-level governance

Public sector



private sector



operational plan

3. Spatial Analysis of Financial Data

Extraction of asset geolocations from sustainability reports

Sustainability Report 2023

Water continued

2023 performance

In 2023, we withdrew 950 million m³ of water (2022 restated: 1,080 million m³). The decrease is primarily related to less rainfall year-on-year at some of our Australian industrial coal assets, as well as the sale, closure or placing on care and maintenance of some industrial sites.

We aim to have no serious environmental incidents involving water at our industrial assets, and to prevent unplanned water discharges. Under our environmental licensing conditions, some of our industrial assets are permitted to discharge surplus water, subject to varying conditions.

We investigate significant environmental incidents and, where appropriate, work with local authorities to address any operational impacts on water sources. In 2023, we recorded seven 'moderate' environmental incidents (2022: three 'moderate' incidents – see 'incident classification' in the 2023 Group Reporting Glossary for classification details). Of these, two that impacted the land are detailed in the Nature section on page 48. Immediate corrective measures were taken, the incidents

reported, and investigations were undertaken in line with regulatory requirements.

The incidents impacting water included:

- As a result of an extreme rainfall event, there was a stormwater release from a spillway at Queensland Metals' Lady Loretta zinc mine in Australia into the Desert Creek catchment, temporarily affecting water quality. Immediate corrective actions were implemented to minimise the impact and additional preventative measures were identified for action through an incident investigation. Potentially affected persons downstream of Desert Creek catchment were contacted during the release.
- An overflow of process water at Volcan's Carahuacra zinc mine in Peru temporarily affected the water quality of the Yauli River and the soil of the adjacent industrial area. Immediate actions were taken to minimise the impact and to remove sediment from the impacted area.
- A diesel spill from a pipeline switch station associated with our Astron Energy oil refinery in South Africa impacted the immediate area and the down-gradient

GLENCORE

Energising today
Advancing tomorrow

2023 Group water balance (ML)

Water input	949,715
Surface water	375,509
Sea water	159,800
Groundwater	333,506
Third-party water	80,906

Operational water use	596,390
Change in water in storage	-93,288

Reused/recycled 501,073

Other managed water – withdrawn 9,712

<https://www.glencore.com/.rest/api/v1/documents/static/be5b0554-2c1d-415d-8072-be6a30d91d79/GLEN-2023-Sustainability-Report.pdf>

ACLED

Clarity in Crisis

Implementación del programa de seguridad "Volcán seguro" para la minería de cobre en la zona de influencia de la mina de Carahuacra, Perú

Spatial Cues in Corporate Sustainability Reports: the Sample

The sample was obtained by combining three curated samples of

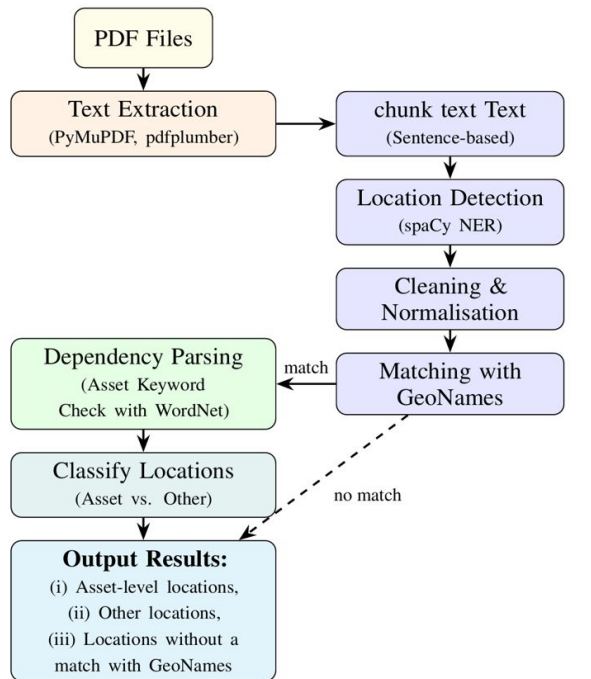
- Corporate Sustainability Reporting Directive (CSRD)-compliant reports compiled by the Sustainability-Reporting-Navigator research project (nreports = 324)
- and the data provider Key ESG (nreports = 148)
- sample of reports from a repository maintained by the Deutsche Bundesbank data scientists (nreports = 368).

Variable	Description
Report Type	Annual / Integrated report: 577 Sustainability report: 263
Report Pages	Min.: 12.0 1st Qu.: 109.0 Median: 208.0 Mean: 242.3 3rd Qu.: 334.2 Max.: 1183.0
Report Year	2023: 405 2024: 435
Report Language	DE: 7 EN: 831 FI: 1 NL: 1

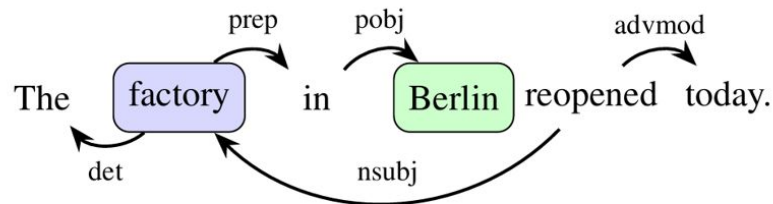
Source: Dr. Felicitas Sommer, Dr. Andreas Dimmelmeier, Ahmed Elhefnawi, Marcel Metzger, Christophe Christiaen, Dr. Emily Kormanyos (forthcoming): Spatial Cues in Corporate ESG Disclosure: An Interdisciplinary Research Agenda for NLP-Driven Analysis. In: Finance and Space.

Spatial Cues in Corporate Sustainability Reports: the sample

Process flow of the asset location extraction pipeline



Dependency parsing process chart



det
prep
pobj
nsubj
advmod

determiner (e.g., "the")
preposition introducing a phrase
object of a preposition
nominal subject (who or what acts)
adverbial modifier (e.g., "today")

asset keyword

location (GPE)

Source: Dr. Felicitas Sommer, Dr. Andreas Dimmelmeier, Ahmed Elhefnawi, Marcel Metzger, Christophe Christiaen, Dr. Emily Kormanyos (forthcoming): Spatial Cues in Corporate ESG Disclosure: An Interdisciplinary Research Agenda for NLP-Driven Analysis. In: Finance and Space.

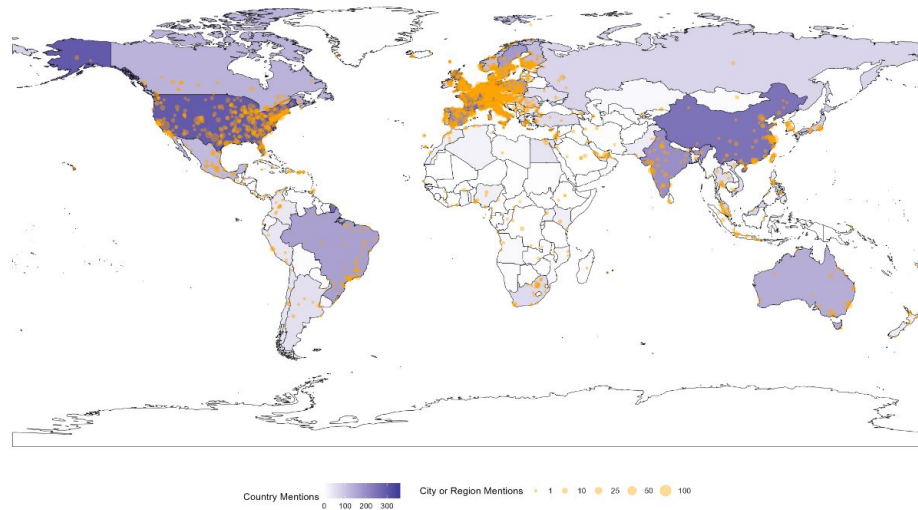
Spatial Cues in Corporate Sustainability Reports: the result

Q1: geographical distribution aligns fairly with sample distribution

Q2: space intensive sectors (example: storage and transportation) report more spatial information than non-space intensive sectors (information and communication)

Q3: sustainability information (which is the biggest topic for location mentions) is more connected to governance topics than to "pure" asset information

Map with extracted country mentions (n = 137) and geocoded extracted cities or regions (n = 2579)

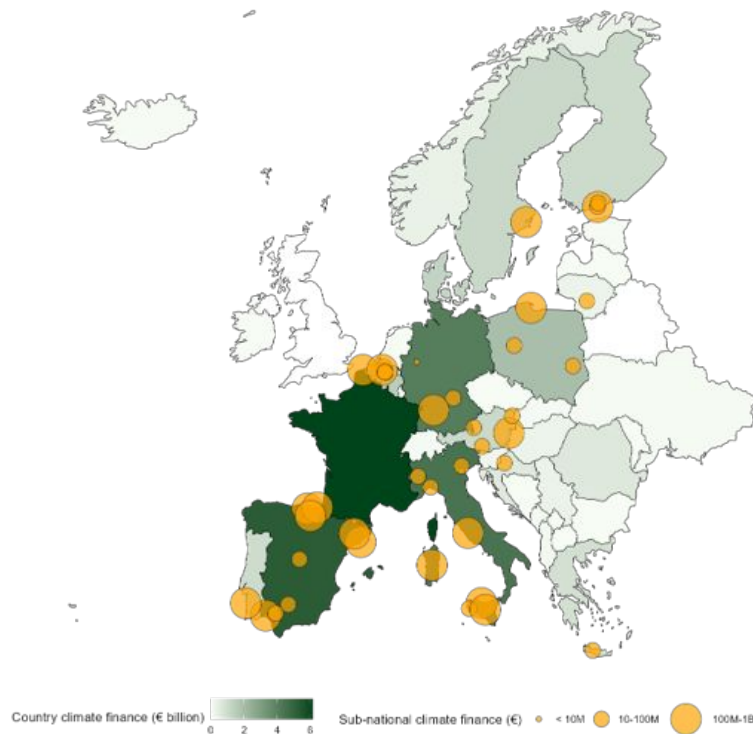


Source: Dr. Felicitas Sommer, Dr. Andreas Dimmelmeier, Ahmed Elhefnawi, Marcel Metzger, Christophe Christiaen, Dr. Emily Kormanyos (forthcoming): Spatial Cues in Corporate ESG Disclosure: An Interdisciplinary Research Agenda for NLP-Driven Analysis. In: Finance and Space.

Comparing Stoxx600 sustainability reports and EIB CA&ES loanbook

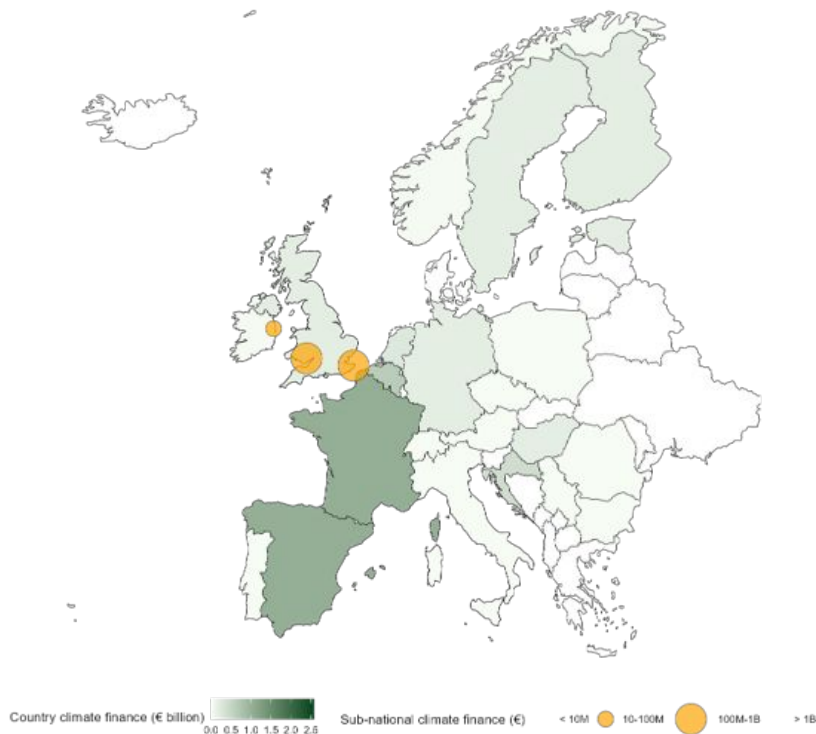
EIB climate finance in Europe

Climate Action amount (€), 2023-2024. Source: EIB CA&ES detailed lists joined to loanbook.



Stoxx 600 finance: climate investments in Europe

Regex-tagged climate finance, 2023-2024 sustainability reports.



How to move forward with the climate adaptation data gap?

Obstacles towards accessible, reliable, coherent, reusable sustainable finance data

- Lack of geospatial data
- Intransparent methodologies of commercial data providers and ‘aggregate confusion’ limit utility of corporate sustainability data
- Limitations of high-quality AI-supported information extraction
- Delayed introduction of central access point and structured machine-readable corporate sustainability reports in the EU

Addressing Data Gaps in Climate Change and Sustainability Statistics using Generative AI¹

Maurice Fehr¹, Simon Oehler¹, Lisa Reichenbach¹, Marvin Riemer¹, Anna Steinberg Schulten², Malte Schierholz², Andreas Dimmelmeier², Felicitas Sommer³

Abstract

Climate change has emerged as a critical macroeconomic and financial risk. Effective policy responses require high-quality, granular sustainability data - in particular issuer-level greenhouse gas (GHG) emissions - which remain scarce, inconsistent, and difficult to extract from corporate disclosures. This paper presents the “Greenhouse Gas Insights and Sustainability Tracking” (GIST) project, a collaboration between Ludwig-Maximilians-University Munich and Deutsche Bundesbank, which leverages state-of-the-art generative artificial intelligence to address these data gaps. We introduce a Retrieval-Augmented Generation (RAG) pipeline for extracting Scope 1–3 GHG emissions from corporate reports, converting unstructured disclosures into standardized, analysis-ready datasets. The pipeline is validated against a rigorously constructed gold standard dataset involving both non-expert and expert review. Results demonstrate that large language models (LLMs) achieve high extraction accuracy (precision: 0.78; recall: 0.90), though challenges persist in handling ambiguous or fragmented disclosures. We further discuss the development of an open-source Python package enabling broad adoption and customization of the extraction workflow. Real-world applications within Deutsche Bundesbank highlight the pipeline’s potential to enhance data quality, reduce manual labor, and support evidence-based policy making.

Keywords: Data gaps, sustainable finance data, climate data, sustainability reports, natural language processing, large language models

JEL classification: E58, Q56, C55, C63, C81, C88

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

¹We would like to express our sincere gratitude to the BIS Innovation Hub Eurosystem Centre (Project GAIA) and the Deutsche Bundesbank Directorate General Markets (ICAS) for their support and fruitful exchanges throughout the course of this project.

²Deutsche Bundesbank

³Ludwigs-Maximilians-Universität München

⁴Technische Universität München

Forthcoming publications with co-authors Andreas Dimmelmeier and Felicitas Sommer



UNITED NATIONS ECONOMIC COMMISSION FOR EUROPE

CONFERENCE OF EUROPEAN STATISTICIANS

Expert meeting on Statistical Data Editing

19 to 21 October 2026, Wiesbaden



Technical Fixers or Epistemic Stakeholders? Data Editing in Wealth and Emissions Statistics

Andreas Dimmelmeier (Ludwig-Maximilians-Universität München, Germany), Pirmin Fessler (Österreichische Nationalbank, Austria), Felicitas Sommer (Ludwig-Maximilians-Universität München, Germany)

a.dimmelmeier@lmu.de; pirmin.fessler@onb.at; felicitas.sommer@lmu.de

I. The expanded role of editing in the re-design of data products

How is a new statistical artefact created? The UNECE’s General Statistical Business Process Model (GSBPM) answers this question with a codified chain where societal information needs are translated into measurable concepts in a design phase, which are then operationalized through data collection and data processing. Afterwards the created statistical products are analyzed, disseminated and, finally, evaluated. Should the evaluation find new needs or required changes, a feedback loop to the beginning of the chain starts the process anew.

While the GSBPM explicitly states that this procedural chain is not rigid and acknowledges the real-world practice of recursive operations in the later stages [UNECE, 2025, p. 5], the model description highlights that the first phases concerning the identification of needs and their translation through design choices are only rarely part of such recursive changes in National Statistical Offices (NSOs) (ibid.).

In this paper, we examine how new information needs are accommodated within existing statistical infrastructures through data editing. Rather than always leading to an explicit redesign of concepts and collection instruments, changing requirements may be addressed through decisions about which observations to retain, correct or impute. Such decisions can help constitute the statistical objects subsequently used in research and policy. This does not imply that feedback to earlier stages disappears, but that editing may accommodate demands that have not been fully addressed through design.

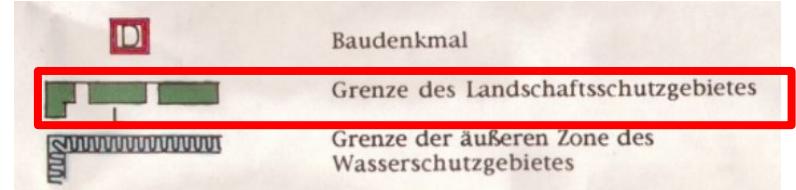
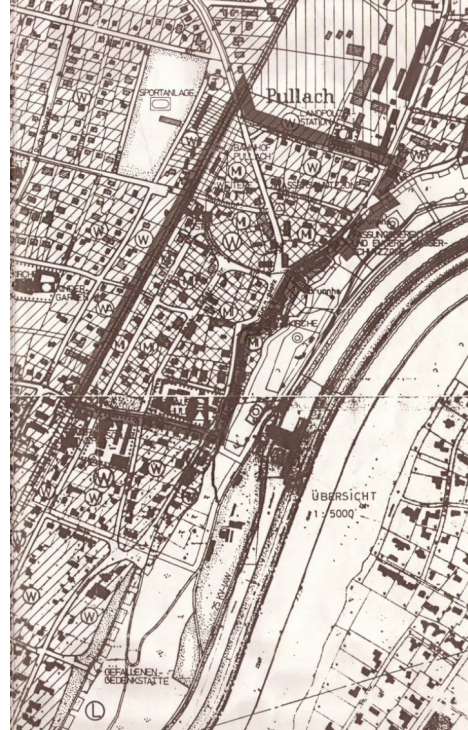
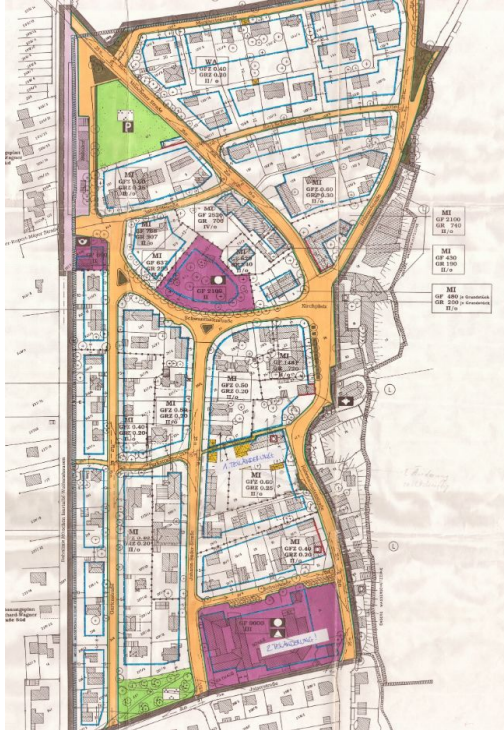
This creates a tension: the more successfully editing makes incomplete or ambiguous data usable, the less visible the original limitations may become to downstream users. Editing can reduce some uncertainties while introducing or transforming others, including dependencies on production choices and unresolved questions about conceptual boundaries. We therefore examine statistical, production and conceptual uncertainty as related dimensions of editing, while also considering how editing stabilizes statistical objects and enables their use in new contexts.

To develop this account, we present two distinct cases. The first concerns the Household Finance and Consumption Survey (HIFCS), a harmonized household survey coordinated within the European System of Central Banks (ESCB) and conducted by national central banks and statistical institutes. Its uses in monetary policy, financial stability and wealth inequality analysis place different demands on the same data infrastructure. These demands become particularly visible in the treatment of missing

4. Regulatory information from land-use plans

Land-use Planning in Germany

Planning documents indicate the will of municipalities and detailed environmental, social, legal conditions



Für alle Neubauten und Umbauten gilt für die lärmbeaufschlagten Fassaden entlang der Bahnlinie bzw. der Münchener Straße folgende Festlegung:

Aufenthaltsräume sind durch technische Vorkehrungen (z.B. Schallschutzfenster) so gegen Aussenlärm zu schützen, daß einschließlich etwaiger Belüftungseinrichtungen oder Belüftungen von Schlaf- und Kinderzimmern über Wintergärten, am Tag ein Innenschallpegel von 30 dB(A) und bei Nacht ein solcher von 25 dB(A) nicht überschritten wird (VDI 2719).

Wird vor notwendigen Fenstern von Schlaf- und Kinderzimmern nachts ein Beurteilungspegel von 45 dB(A) überschritten, ist für eine ausreichende Belüftung der Räume über vorgebaute Wintergärten oder schallgedämpfte Belüftungseinrichtungen Sorge zu tragen.

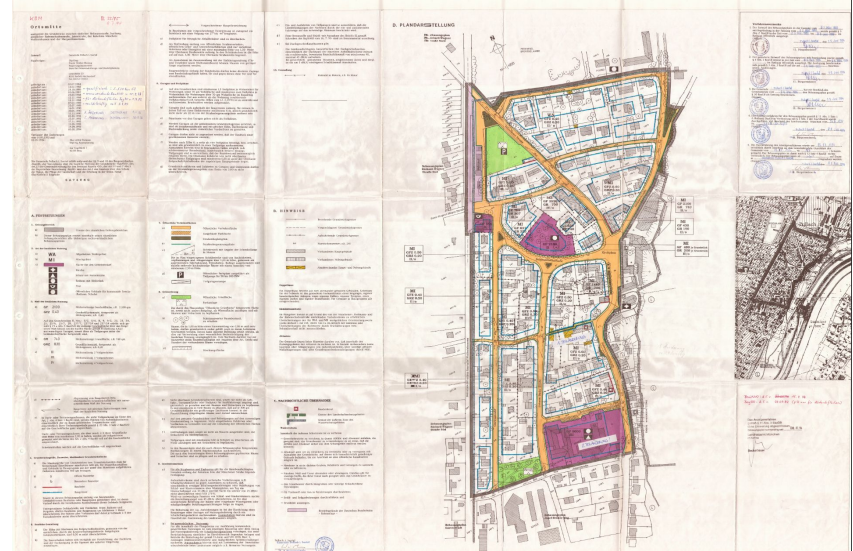
Nicht überbaute Grundstücksflächen sind, soweit sie nicht als Geh-, Fahr-, Terrassenfläche oder Stellplatz für Kraftfahrzeuge angelegt sind, gärtnerisch zu gestalten und mit Bäumen und Sträuchern zu bepflanzen. Es sind mindestens so viele Bäume zu pflanzen, daß auf je 300 qm Grundstücksfläche ein großkroniger Laubbaum kommt; in der Planzeichnung eingezeichnete Räume sind darauf anzuordnen.

Land-use Planning in Germany

Accessibility of planning documents is 'built into' information structure, document formats, and scan quality

Challenges for automated information extraction from binding land-use plans:

- Multi-modality (text/image)
- Scan quality (earliest plans from 1936)
- Low standardization of formats and symbols (despite planning symbol ordinance)
- accessibility' of spatial planning documents

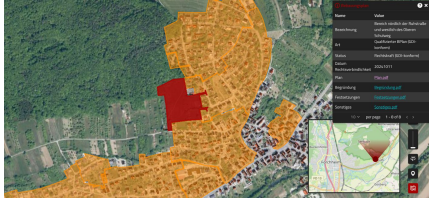


Source:

https://wms1.rivadis.de/muenchen_lk_bp/bp-dateien/pullach%20i.isartal/a7-713-0022-85-bl%20bebauungsplan.pdf

Building high/density from land-use plans

1 Municipal Portal

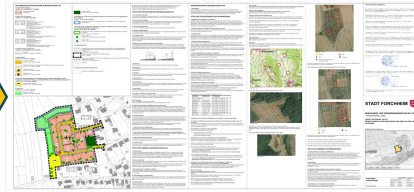


Shapefile Attribute

Select the plan polygon and retrieve its linked documents from the shapefile attributes.

Example: Reuth, Oberer Schulweg - Bebauungsplan 10/6

2 Document



MapURL + TextURL (where available)

Pdfs Plan & Reasoning

Download the scan and text; keep the plan ID and document type with each file.

3 Optical Character Recognition (Paddle)



Texts and location

Clean overlapping tile fragments while preserving genuine repeated occurrences.
"GRZ 0,4"
[980, 1130, 1162, 1180] px

4 Large Language Model Gemma 4

SCAN · page 1

GRZ 0,4

Reasoning · page 33

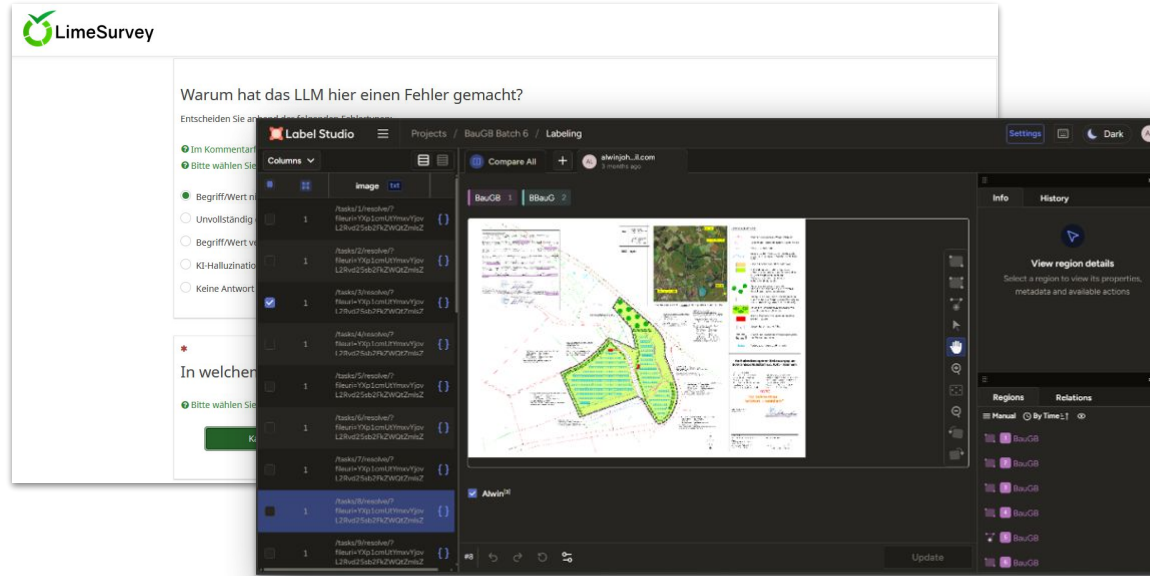
GRZ max. 0,4,
max. 0,6 for terraced middle houses on small plots
GFZ: no extraction in either

Structured Extraction

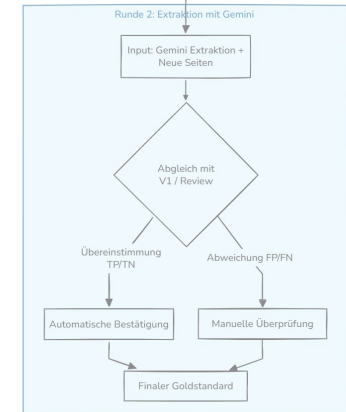
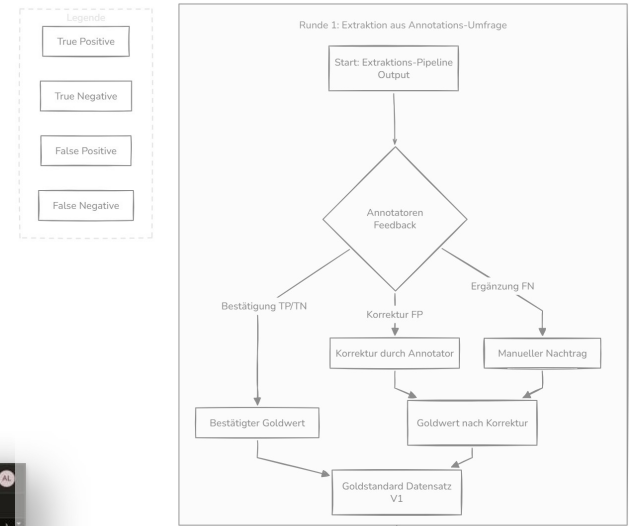
Apply the OCR prompt; retain metric, value, qualifiers and source evidence in JSON.

Challenges of AI Extraction

Human-in-the-Loop-Approach

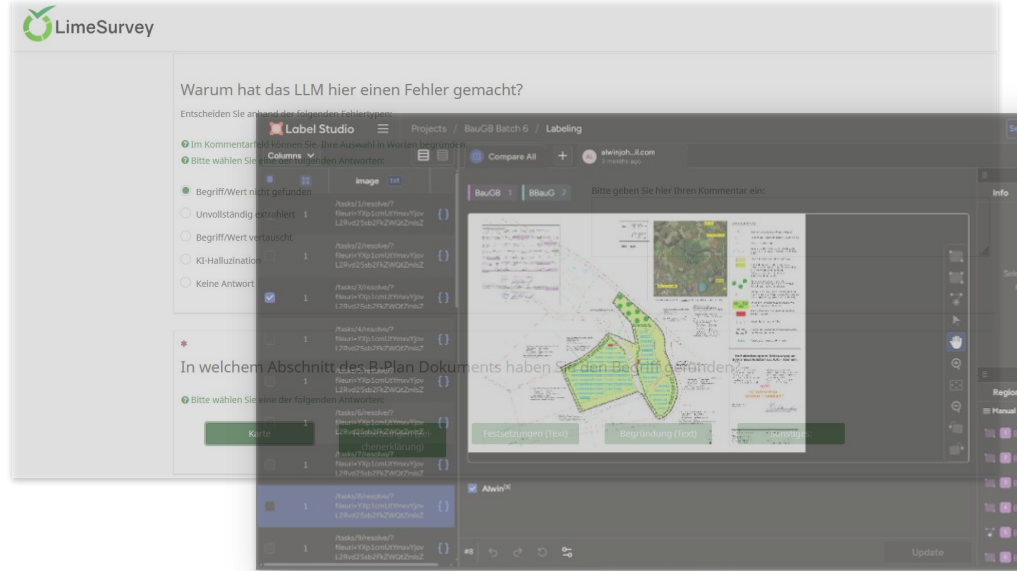


Human Labeling

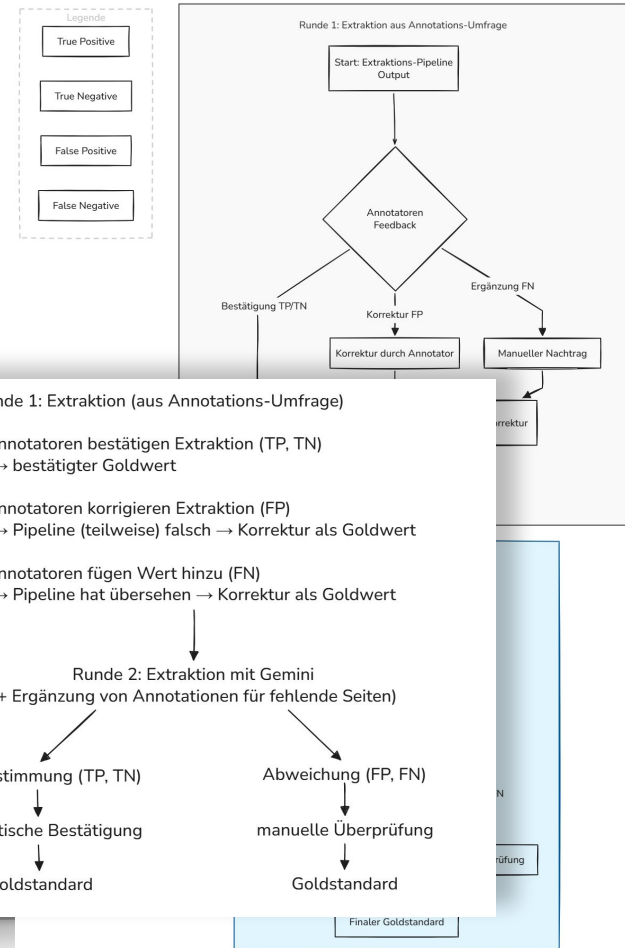


Comparison Human-AI

Human-in-the-Loop-Approach



Human Labeling



Comparison Human-AI

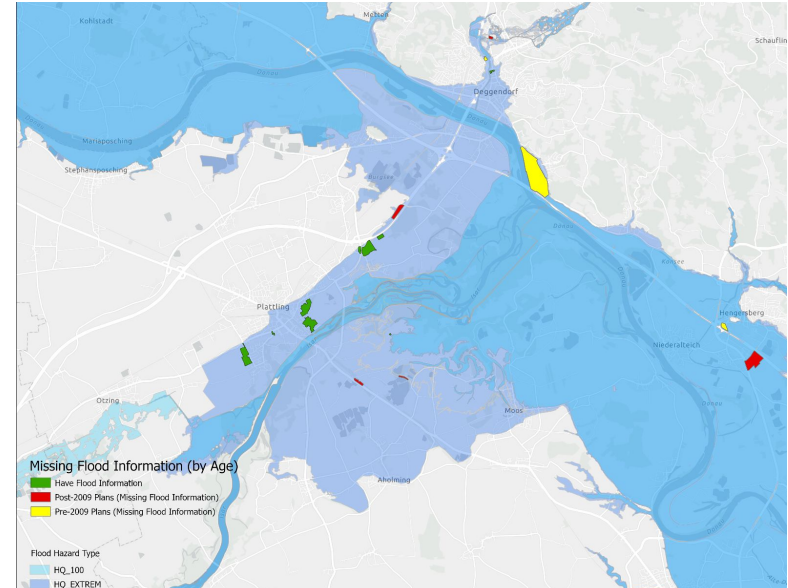
Links between flood management and Land-use Planning

Machine-readable digital land-use plan data can be used to assess how effective new climate adaptation regulation has been implemented in land-use planning.

Example:

How did land-use planning respond to the 2009 amendment of the Water Resources Act implementing the EU Floods Directive ?

Land-use plans with or without flood risk information before and after 2009



Sources: Map Flood Hazard Risks: LfU; Land-use Plans: Bayernatlas & B-Planpool.
Authors: Alwin John Mathews, Sebastian Riznik, Felicitas Sommer.

More or less flood risk management information after 2009?

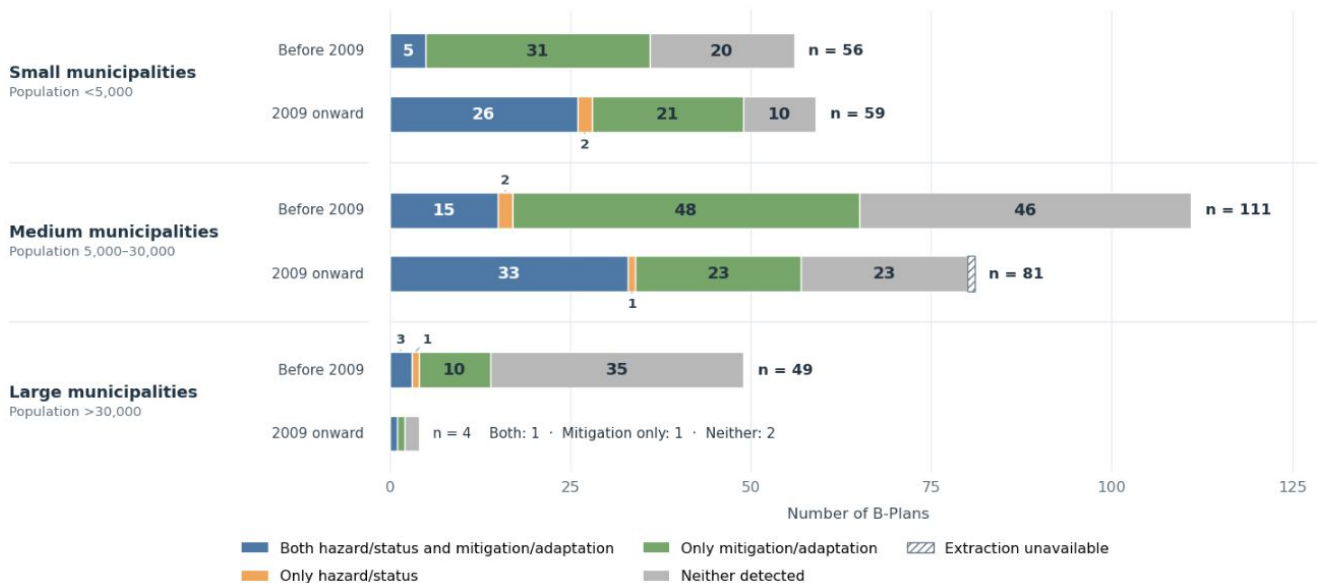
Small and medium sized municipalities seem to have **more** flood risk information than large ones after 2009.

Note: the sample contains large municipalities less (only 53 of 360 plans)

Planning content by municipality size and plan period

Preliminary comparison of pre-2009 and post-2009 B-Plans

Bar length = total plans | Segment labels = number of plans

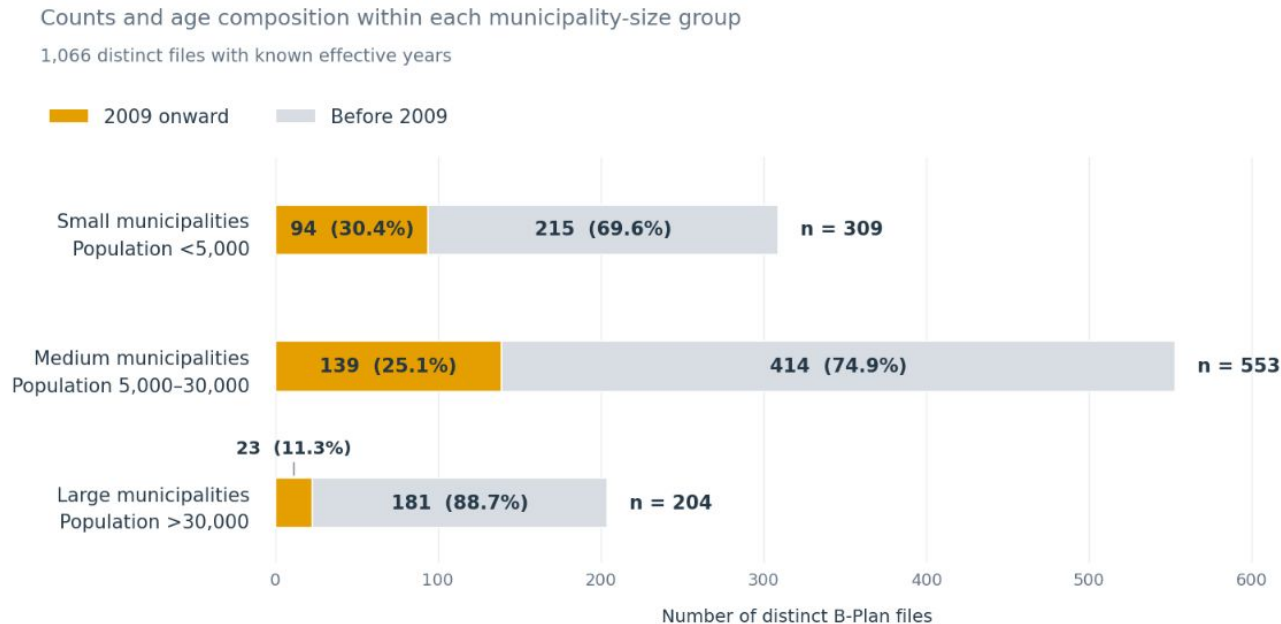


Neither = neither selected content dimension detected. Medium municipalities, 2009 onward: 1 extraction unavailable.

Preliminary: subgroup sizes are uneven and results should not be interpreted as a ranking of municipal performance.

Large municipalities stop building in flood risk areas!

Land-use plans in larger municipalities contain less flood risk information - because they **reduced development in high-risk areas** after 2009.



Bar length shows total files. Segment labels show count and percentage within each municipality-size group.

One record per filename; unknown years excluded. 2009 belongs to the later period. Population reference: 30 September 2025.

Preliminary: available-file counts do not measure annual planning rates, policy effects or municipal performance.

5. Land management information and federated data architectures

Integrated Land Management from a Systemic Perspective

What is the traditional domain of environmental risk mitigation?

Land-use planning: is cross-sectoral, multi-level, multi-functional

Public restrictions and private rights: are transparent, create accountability, enable implementation and enforcement

Soil-function maps: consider all-hazards, green infrastructures, co-benefits, allow feedback loops

Buildings and infrastructure registers: allow stocktaking of demands and capacities, evidence-based decision-making

Integrated Land Management from a Systemic Perspective

What is the traditional domain of environmental risk mitigation?

➡ Use Spatial Data Infrastructures (SDIs) and Land Information Systems for climate adaptation information?

Land-use planning: is cross-sectoral, multi-level, multi-functional

Public restrictions and private rights: are transparent, create accountability, enable implementation and enforcement

Soil-function maps: consider all-hazards, green infrastructures, co-benefits, allow feedback loops

Buildings and infrastructure registers: allow stocktaking of demands and capacities, evidence-based decision-making

Advancing the Land Management Paradigm

How can land management systems inform fiscal decision making or prudential regulation and administrative architectures?

- Land registers can enrich corporate asset ownership information
- Spatial planning information can corroborate alignment with climate adaptation planning and regulation

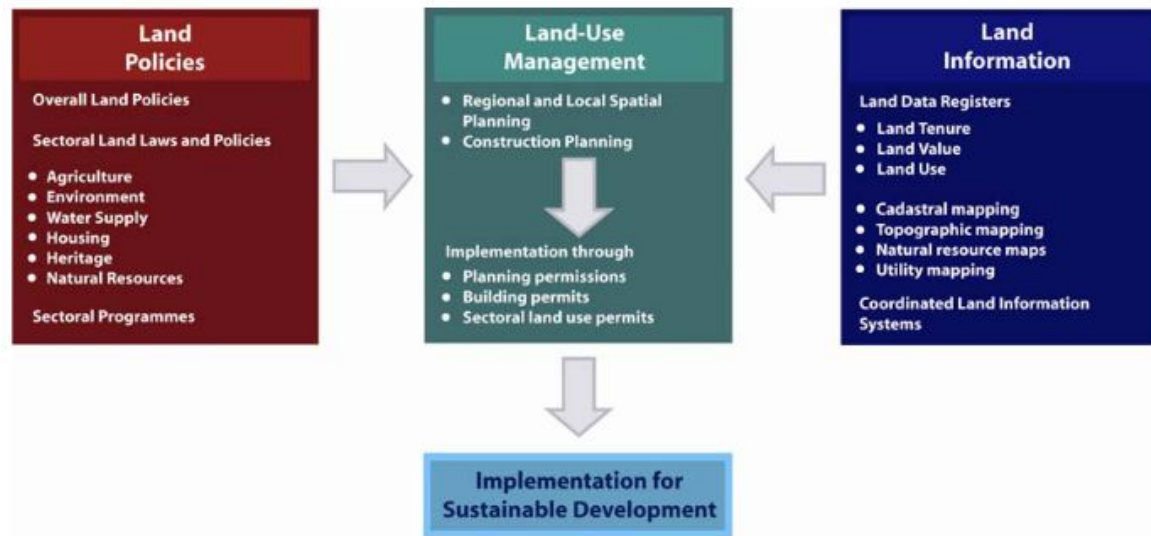


Figure 5. Integrated Land-Use Management for Sustainable Development.

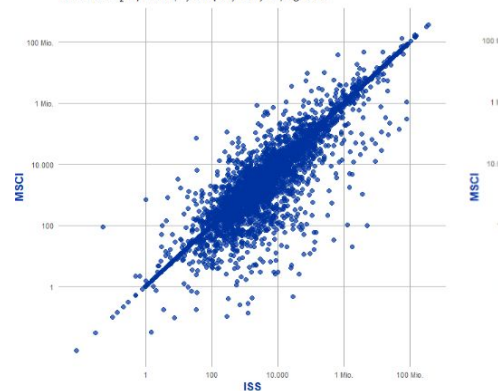
Conclusion

There is

- no time for 'aggregate' confusion and consequential assumption errors
- no capacity for siloed, fragmented, redundant data production
- no space for ineffective adaptation and conflicting land-use activities

➡ Every step towards more federated, cross-sectoral, all-hazard, multi-level data architectures is a step towards climate resilience

Comparison of scope 1 GHG emissions
in tons of CO₂ equivalent, by company and year, log scale



Source: Fehr et al.: Addressing Data Gaps in Climate Change and Sustainability Statistics using Generative AI. @ 25th BIS Annual Conference



Source:
<https://www.vallartadaily.com/vallarta-storm-damage-raises-questions-over-coastal-permits/>

Thank you for your attention!

Felicitas.sommer@lmu.de

Felicitas.sommer@tum.de

References

- Aerts, J.C.J.H. (2018) "A Review of Cost Estimates for Flood Adaptation," *Water*, 10(11), p. 1646. Available at: <https://doi.org/10.3390/w10111646>.
- Aerts, J.C.J.H. et al. (2024) "Exploring the limits and gaps of flood adaptation," *Nature Water*, 2(8), pp. 719–728. Available at: <https://doi.org/10.1038/s44221-024-00274-x>.
- Andrews, T.M. et al. (2023) "Risk from responses to a changing climate," *Climate Risk Management*, 39, p. 100487. Available at: <https://doi.org/10.1016/j.crm.2023.100487>.
- Condon, M. (2023) "Climate Services: The Business of Physical Risk," *Arizona State Law Journal*, 55(1), p. 147.
- De Plaen, J.J.-F.G. et al. (2025) "COASTPROS-EU." Available at: <https://doi.org/10.5281/zenodo.15024139>.
- Di Baldassarre, G. et al. (2017) "Drought and flood in the Anthropocene: feedback mechanisms in reservoir operation," *Earth System Dynamics*, 8(1), pp. 225–233. Available at: <https://doi.org/10.5194/esd-8-225-2017>.
- Directorate-General for Financial Stability, F.S. and C.M.U. et al. (2024) "Methodological Framework for Assessing Financial Risks from Biodiversity Loss and Ecosystem Degradation." Available at: <https://policycommons.net/artifacts/11770868/study-for-a-methodological-framework-and-assessment-of-potential-financial-risks-associated-with-biodiversity-loss-and-ecosystem-degradation-final-report/12664327/> (Accessed: October 18, 2026).
- Dubou, G., Bichueti, R.S. and Kneipp, J.M. (2026) "Mapping urban resilience strategies to climate change: a systematic review on absorption, adaptation, and transformation capacities," *Environment, Development and Sustainability* [Preprint]. Available at: <https://doi.org/10.1007/s10668-026-07952-8>.
- Enemark, S. (no date) "Understanding the Land Management Paradigm."
- European Central Bank (2024) "Analytical indicators on physical risks." Available at: https://www.ecb.europa.eu/stats/all-key-statistics/horizontal-indicators/sustainability-indicators/data/html/ecb.climate_indicators_physical_risks.en.html (Accessed: October 16, 2026).
- European Central Bank. (2024) Climate change-related statistical indicators. LU: Publications Office. Available at: <https://doi.org/10.2866/059096>.
- van Maanen, N. et al. (2025) "Brief communication: Bridging the data gap – a call to enhance the representation of global coastal flood protection," *Natural Hazards and Earth System Sciences*, 25(6), pp. 2075–2080. Available at: <https://doi.org/10.5194/nhess-25-2075-2025>.
- OECD (2025) Government at a Glance 2025. OECD Publishing (Government at a Glance). Available at: <https://doi.org/10.1787/0efd0bcd-en>.
- Rohmer, J. et al. (2026) "Uncertainty in broad-scale coastal flood damage cost assessments under sea-level rise: critical needs for more accurate protection data," *Environmental Research Communications*, 8(10), p. 105001. Available at: <https://doi.org/10.1088/2515-7620/aea67b>.
- Schröter, K. et al. (2025) "Invited perspectives: Fostering interoperability of data, models, communication, and governance for disaster resilience through transdisciplinary knowledge co-production," *Natural Hazards and Earth System Sciences*, 25(9), pp. 3055–3073. Available at: <https://doi.org/10.5194/nhess-25-3055-2025>.
- Simpson, N.P. et al. (2025) "Advances in complex climate change risk assessment for adaptation," *npj Climate Action*, 4(1), p. 74. Available at: <https://doi.org/10.1038/s44168-025-00281-y>.
- Walter, S. (2025) "Estimating companies' GHG emissions using information from individual financial accounts and building data," *Statistical Journal of the IAOS*, 41(2), pp. 305–316. Available at: <https://doi.org/10.1177/18747655251341456>.
- Bharwani, S., Williamson, K. and Wangrat, A. (2026) "Scaling knowledge interoperability for accelerated climate action through the Climate Connectivity Hub and Taxonomy," *Climatic Change*, 179(3), p. 48. Available at: <https://doi.org/10.1007/s10584-026-04139-y>.
- Fouilloux, A. et al. (2026) "FAIR Data Infrastructure: The Missing Layer in EU Climate Resilience Policy." Available at: <https://doi.org/10.5281/zenodo.18743694>.

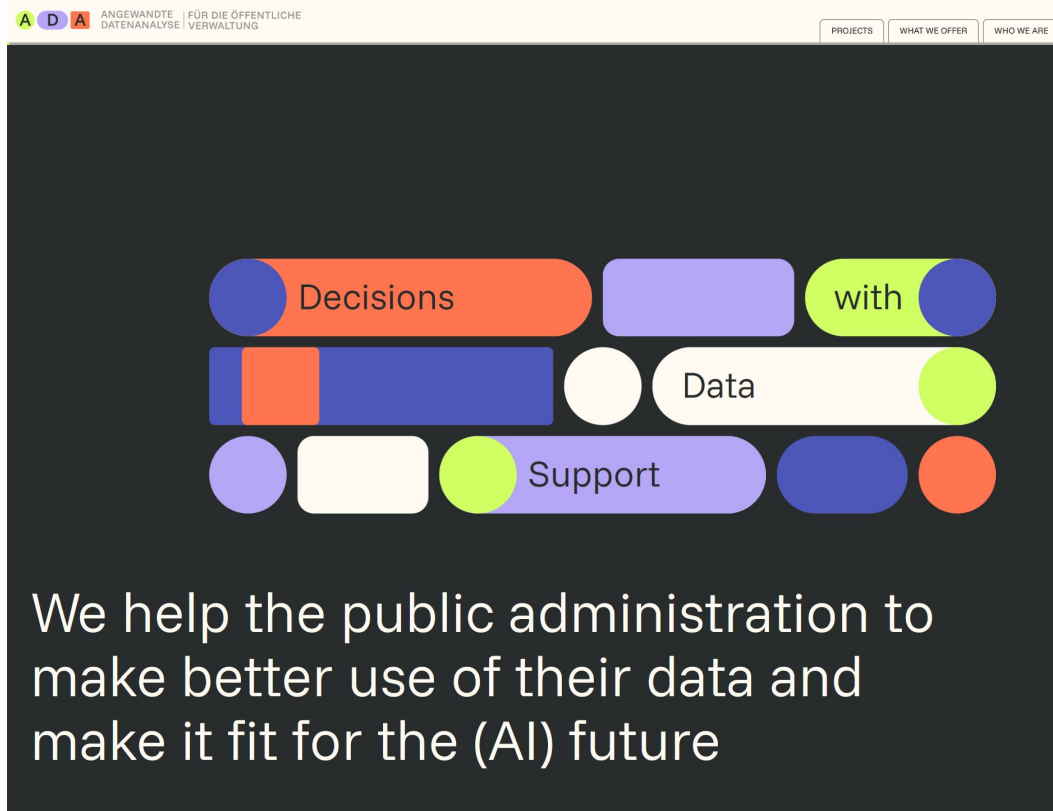
Is cross-sectoral governance information
integration possible?

Evidence-based decision making: data literacy for administrations

ADA (Applied Data Analysis for public Administrations) supports digitization and digital literacy.

The aim of enabling and training public administrations is to accelerate effective, machine-readable, harmonized data architectures for evidence-based decision making.

The digital transition shouldn't just replicate old systems but integrate current informational and operational demands.



The need for FAIR Climate Adaptation Knowledge

- FAIR Data Infrastructure: The Missing Layer in EU Climate Resilience Policy (Fouilloux et al. 2026)
- knowledge interoperability for accelerated climate action through the Climate Connectivity Hub and Taxonomy (Bharwani et al. 2026)
- fostering interoperability of data, models, communication, and governance for disaster resilience through transdisciplinary knowledge co-production (Schröter et al. 2025)



eosc | FAIR2Adapt

FAIR2Adapt

Advancing climate
adaptation in Europe

Making climate data easier
to find, share, and reuse, so
regions and cities across
Europe can build stronger,
more resilient futures.

Microdata Infrastructure

New database developed by Austrian National Bank

Challenge: combined datasets had incompatible identifiers, heterogeneous variable names, divergent temporal structures and refer to different and often implicit or undefined target populations

Potential: prevent error-prone preprocessing

OeNB Report 2026/27
46 pages

The Integrated Firm-Level Database (IFLD): a new microdata infrastructure for analysing Austrian firms

Structure and first empirical insights for 2021

The Integrated Firm-Level Database (IFLD) combines administrative and supervisory data on Austrian firms into a unified microdata infrastructure. Its key innovation is that it links previously fragmented information on firm characteristics, financial statements, bank lending, ownership and insolvencies within clearly defined firm populations. The IFLD reveals that the Austrian economy is dominated in number by small firms, while economic activity and financing are highly concentrated among relatively few firms. Ownership is often concentrated among a small number of owners, predominantly natural persons.

JEL classification
C81, D22, G21, G32, L25

Keywords
Integrated firm-level database, corporate sector, bank lending, bank credit, granular credit data, firm financing, ownership networks, corporate insolvency, microdata, Austria

Authors

Helmut Elsinger, Pirmin Fessler, Alexandra Riedl, Philipp Rodewies, Stefan Trapp
Oesterreichische Nationalbank
Research

Laura Fehring, Bernhard Fuchs, Reinhard Isler, Thomas Krametschler, Barbara Riedrich, Lukas Simhandl, Stefan Weisinger
Oesterreichische Nationalbank
Statistics Department

Publication
8 October 2026

Corresponding authors
pirmin.fessler@oebn.at
alexandra.riedl@oebn.at



Small firms dominate by number

The majority of Austrian non-financial firms are very small. Among firms subject to disclosure requirements for which information is available, the median firm has two employees, an annual turnover of around EUR 369,000 and total assets of around EUR 547,000. A comparatively small number of large firms account for much of economic activity.



Direct bank financing is limited

Only around 36% of Austrian non-financial firms have a loan from an Austrian credit institution. Also, bank lending is highly concentrated, with a small share of borrowers accounting for the bulk of outstanding credit.



Ownership patterns differ sharply by firms' economic weight

In 62% of limited liability companies, the majority shareholder is a single natural person. At the same time, firms majority-owned by a single person account for 79% of total assets. This shows that corporate ownership dominates among economically important firms.

Approaches for spatial finance with NLP and remote sensing

- Universal ownership is a “harmful fiction” (Prodani 2025)
- Geospatial Finance links economic and financial attributes to environmental and spatial resources (Gopal and Pitts, 2024)
- Multimodal Industry Classification through Geographic Information with Multi Agent Systems can accelerate identification of assets (see image, Yüksel et al. 2025)

MONETA: Multimodal Industry Classification through Geographic Information with Multi Agent Systems¹

Arda Yüksel^{1,2}, Gabriel Thiem^{3,4}, Susanne Walter³, Patrick Felka¹, Gabriela Alves Werb^{1,4}, Ivan Habernal^{1,4,5}

¹Trustworthy Human Language Technologies ²Technical University of Darmstadt, Germany

³Deutsche Bundesbank ⁴Frankfurt University of Applied Sciences, Germany

⁵Research Center for Trustworthy Data Science and Security, Ruhr University Bochum, Germany

Correspondence: arda.yueksel@tu-darmstadt.de

Abstract

Industry classification schemes are integral parts of public and corporate databases as they classify businesses based on economic activity. Due to the size of the company registers, manual annotation is costly, and fine-tuning models with every update in industry classification schemes requires significant data collection. We replicate the manual expert verification by using existing or easily retrievable multimodal resources for industry classification. We present MONETA, the first multimodal industry classification benchmark with text (Website, Wikipedia, Wikidata) and geospatial sources (OpenStreetMap and satellite imagery). Our dataset enlists 1,000 businesses in Europe with 20 economic activity labels according to EU guidelines (NACE). Our training-free baseline reaches 62.10% and 74.10% with open and closed-source Multimodal Large Language Models (MLLM). We observe an increase of up to 22.80% with the combination of multi-sum design, context enrichment, and classification explanations. We release MONETA and code¹.

1 Introduction

Geospatial finance (Gopal and Pitts, 2024) is an emergent and complex field that links financial and economic attributes to environmental and spatial resources. One of the important milestones in multimodal AI research, Multimodal Large Language Models (MLLM) such as Llava (Jiu et al., 2023b,a, 2024), InternVL (Chen et al., 2024b; Wang et al., 2024b), QwenVL (Bai et al., 2023; Wang et al., 2024a; Bai et al., 2025), GPT-5 (OpenAI, 2025) and Gemini 2.5 (Gemini 2.5 Team, 2025) can contribute to geospatial finance decision making tasks by processing visual geospatial data in addition to text documents. Traditionally, AI research has developed unimodal automatic industry

¹The views expressed in this paper are personal views of the authors and do not necessarily reflect the views of Deutsche Bundesbank or the European Central Bank.

²<https://github.com/trustworthy/moneta>

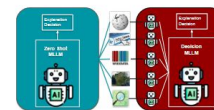


Figure 1: Zero-Shot vs Multi-Turn comparison for MONETA. (1) Zero-shot pipeline (left): Available resources are forwarded into the industry classifier MLLM together. Explanations and classifications are obtained. (2) Multi-Turn pipeline (right): Each resource is processed by separate specialized agents. Intermediate clues from these agents are processed by the decision-making agent, returning explanation and classification.

classification based on global (ISIC (UN, 2008)) and region-specific (NAICS (Ambler and Kristoff, 1998), NACE (European Commission, 2008)) industry classification systems. National Statistical Business Registers classify enterprises by their primary economic activity using these classification schemes. The International Standard Industrial Classification (ISIC (UN, 2008)) is widely used for economic statistics, including production, national accounts, employment, and population studies.

Existing research on automatic industry classification from company recordings, financial reports, and websites (Kühnemann et al., 2020; Béchara et al., 2022; Kizinski et al., 2023; Faria and Seimandi, 2023; Vianourellis et al., 2024; Malashin et al., 2024; Guo et al., 2025; Duroyo et al., 2025) has two main drawbacks. First, these methods rely solely on text, which is often unavailable for newly founded or small firms, whereas geospatial information from business registers can provide useful signals. Second, they fine-tune models that require large datasets and limit them to a single classification scheme.

14790

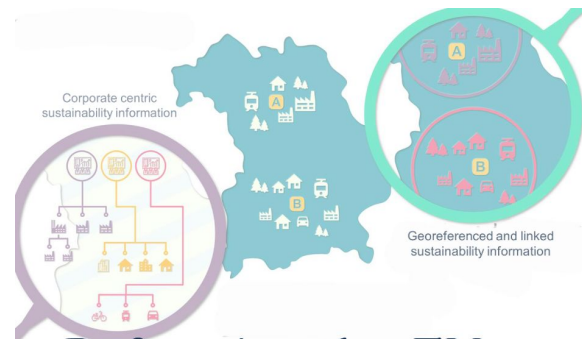
Proceedings of the 64th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers), pages 14790–14814
July 2–7, 2026 ©2026 Association for Computational Linguistics

How to move forward with the climate adaptation data gap?

- Cross-sector coherent land governance matters for reliable (measurement of) state capacity and institutional resilience

Conclusion GreenDIA:

- corporate and financial spatial information is less complete and reliable than geospatial land information -> integrating corporate in geospatial data architectures was more feasible than integrating spatial data in corporate data architectures



Reframing the EU landscape for corporate sustainability information

How moving beyond corporate-centric sustainability information can increase the utility of sustainability data for other key stakeholders critical to the sustainable transition