Digital Sovereignty

Principles 6–7: AI-legibility and third-party anchoring.

- P06 — AI-Legibility
- P07 — Third-Party Anchoring

P06 — AI-Legibility

SCD-P06 · Principle 6 of 10
AI-Legibility
“If an AI agent cannot cite you, you do not exist.”
Digital Layer

Definition

Climate data and impact claims must be structured so that AI agents can discover, parse, cite, and cross-reference them without human intermediation. Minimum requirements: schema.org/Dataset markup on all public pages, JSON-LD structured data, persistent canonical URLs, and inclusion in at least one major open index (Global Forest Watch, Climate TRACE, EDGAR, or Copernicus).

Rationale

The AI+ESG verification market is growing at 28% CAGR and already processes over 100,000 ESG sources daily using NLP models (WEF, 2024). AI agents are used by investors, regulators, and procurement teams for automated due diligence — without notifying the organisations being assessed. An organisation invisible to AI agents is invisible to the decision-makers those agents serve. Critically: AI without grounded sovereign data can also hallucinate plausible-sounding figures — making sovereign data not just a visibility tool but a truth anchor.

Implementation Steps

1. Add schema.org/Dataset JSON-LD to every public data page (see JSON-LD reference below).
2. Submit datasets to Google Dataset Search via schema.org markup.
3. Register with at least one global open data index (CKAN, DataCite, GFW API).
4. Maintain a public llms.txt file (analogous to robots.txt) guiding AI agents to authoritative sources.
5. Run quarterly AI scans: query Climate TRACE, GFW, and EDGAR to confirm your data is indexed.

JSON-LD Reference Example

```
{
  "@context": "https://schema.org",
  "@type": "Dataset",
  "name": "Colombia Deforestation Alerts 2024",
  "description": "GLAD-L primary forest loss alerts for Colombia, 2024.",
  "url": "https://data.cleantechhub.net/datasets/colombia-deforestation-2024",
```

```
  "identifier": "https://doi.org/10.XXXX/cth-col-def-2024",
  "creator": { "@type": "Organization", "name": "CleantechHUB" },
  "datePublished": "2025-01-15",
  "license": "https://creativecommons.org/licenses/by/4.0/",
  "spatialCoverage": { "@type": "Place", "name": "Colombia" },
  "temporalCoverage": "2024-01-01/2024-12-31",
  "keywords": ["deforestation", "Colombia", "GLAD-L", "primary forest", "sovereign data"],
  "distribution": [{
    "@type": "DataDownload",
    "encodingFormat": "application/json",
    "contentUrl": "https://data.cleantechhub.net/api/v1/datasets/colombia-deforestation-2024"
  }]
}
```


Compliance Checklist

	Criterion	What it means
?	**schema.org/Dataset markup live**	JSON-LD is present on all public data pages and passes Google Rich Results test.
?	**Listed in open index**	Dataset appears in at least one: GFW API, Climate TRACE, EDGAR, DataCite.
?	**llms.txt published**	A public llms.txt file at cleantechhub.net/llms.txt guides AI agents.
?	**Quarterly AI scan**	Last scan date recorded; data confirmed indexed in at least one major platform.

Regulatory References

- EU CSRD — Art. 8 (machine-readable XBRL tagging requirement)
- TCFD Recommendations — Pillar 4 (Metrics and Targets, digital disclosure)
- IICSR AI+ESG Market Report 2025

Recommended Tools and Platforms

Google Rich Results Test schema.org validator llms.txt specification DataCite

Keywords

AI legibility schema.org JSON-LD ESG AI llms.txt due diligence NLP

Related Principles: SCD-P01 · SCD-P02
Document ID: SCD-P06 | **Version:** 1.0.0 | **Last Updated:** 2026-05-26 | **Category:** Digital Sovereignty | **Source:** CleantechHUB Sovereign Climate Data Framework | **Licence:** CC-BY 4.0

This page is part of the Sovereign Climate Data Wiki, maintained by CleantechHUB. It is AI-legible, machine-readable, and available via the BookStack REST API.

P07 — Third-Party Anchoring

SCD-P07 · Principle 7 of 10
Third-Party Anchoring
“Your claim is only as strong as its external anchor.”
Digital Layer

Definition

Every sovereign climate claim must be anchored to at least one authoritative external source: a national government dataset (IDEAM, INPE/PRODES, SENAMHI, CONAF), a multilateral intergovernmental platform (GFW, Climate TRACE, EDGAR, Copernicus/ESA), or a certified third-party verifier accredited under an international standard (Verra, Gold Standard, CBAM-accredited VVB). Internal data alone, however well-documented, does not constitute sovereign climate data.

Rationale

Self-reported data without external anchor has zero credibility in investment due diligence, green bond verification, or regulatory audit. Anchoring to a sovereign or intergovernmental dataset transfers the burden of proof from the organisation to the global data infrastructure. For LATAM: IDEAM (Colombia), INPE/PRODES (Brazil), SENAMHI (Peru/Bolivia), and CONAF (Chile) are the authoritative national anchors for forest and climate data.

Implementation Steps

1. Identify the authoritative external source for each reported metric type (forest: GFW/IDEAM; emissions: Climate TRACE/EDGAR; energy: IEA/national grid operator).
2. Cross-reference your data against the external anchor at least quarterly.
3. Document discrepancies between your data and the anchor, with explanation.
4. For carbon credits: work only with verifiers accredited under Verra, Gold Standard, or a national carbon market standard.
5. For CBAM compliance: obtain third-party verification from an EU-accredited VVB.

Compliance Checklist

	Criterion	What it means
?	**External anchor identified per metric**	Each metric type has a named authoritative external source.

	Criterion	What it means
?	**Quarterly cross-reference performed**	Last reconciliation date recorded with discrepancy log.
?	**Verifier accreditation confirmed**	Any third-party verifier's accreditation status has been checked.
?	**CBAM VVB in place (if applicable)**	EU-accredited VVB identified for CBAM-relevant goods.

Regulatory References

- CBAM Regulation EU 2023/956 — Art. 10(3) (Accredited VVB requirement)
- Verra VCS Standard v4.1 — Section 4.2 (Validation and Verification)
- GFW GLAD-L Alert Methodology (Hansen et al. 2013, updated 2024)

Recommended Tools and Platforms

Global Forest Watch API Climate TRACE API EDGAR REST API Copernicus CDS API IDEAM Colombia INPE PRODES

Keywords

third-party verification IDEAM GFW EDGAR CBAM VVB carbon credits Verra

Related Principles: SCD-P01 · SCD-P03
Document ID: SCD-P07 | **Version:** 1.0.0 | **Last Updated:** 2026-05-26 | **Category:** Digital Sovereignty | **Source:** CleantechHUB Sovereign Climate Data Framework | **Licence:** CC-BY 4.0

This page is part of the Sovereign Climate Data Wiki, maintained by CleantechHUB. It is AI-legible, machine-readable, and available via the BookStack REST API.