SOMA AI

Putting your clients' AI use into CSRD Scope 3

A practical brief for sustainability consultants (2026)

Guillermo Llopis-Garcia, CTO, SOMA AI

November 2026

Factor version soma-ai-ef-2026.10-dataset-v3.0.0, factors last updated 19 October 2026. Regulatory position as checked by SOMA on 30 October 2026. This is not legal advice.
somaai.earth

About this brief

This brief is for consultants who are asked to put a client's AI use into a greenhouse gas inventory. It condenses SOMA AI's published guides into one method you can apply and defend.

Every figure comes from SOMA's public pages, with the SOMA AI Emission Factors (SAEF) dataset, version soma-ai-ef-2026.10-dataset-v3.0.0. These are estimates built from published benchmarks and published grid data. SOMA does not meter anyone's data centre.

1. Why it matters now

Most clients are not legally required to report it. Directive (EU) 2026/470, adopted on 24 February 2026, published on 26 February and in force since 18 March 2026, limits the CSRD reporting duty to undertakings with more than 1,000 employees on average and more than EUR 450 million in net turnover. Both conditions apply, for financial years starting on or after 1 January 2027, so reports under the new scope appear from 2028. The Commission estimates about 85% fewer companies in scope. Member States have until 19 March 2027 to transpose these provisions. The thresholds do not depend on AI use, and the Scope 3 requirements in ESRS E1 do not name it.

In Spain, the bill meant to transpose the CSRD (121/000038) is recorded by Congress as lapsed since 6 October 2026, and SOMA found no other Spanish transposing act as of 30 October 2026. The Ley 11/2018 non-financial statement duty is a separate matter.

The questions arrive anyway. Customers, lenders and tender teams send questionnaires, and a company that follows a voluntary standard may want a complete inventory. The Omnibus limits one route. A CSRD reporter cannot require companies in its value chain to provide more information than the voluntary standard covers, and undertakings with up to 1,000 employees get a statutory right to refuse anything beyond that. The data points this covers are listed in Annex II of Regulation 2026/1560, and the Scope 3 paragraphs are not among them, so a protected supplier may decline a request for its own AI line for that purpose. Many will still answer, for commercial reasons or because a lender or tender asks outside that route.

Where the line sits. Purchased AI is Scope 3 Category 1, purchased goods and services. An AI API, an enterprise chat seat and a SaaS tool with AI features are all services the client buys. They are not Scope 2, because the provider buys the electricity, not the client. For a client in CSRD scope, revised ESRS E1 AR 24(g) asks, if applicable, for emissions from purchased cloud computing and data centre services as a subset of Category 1. The text does not name AI, so reading AI services into that subset is SOMA's interpretation. Whether the amount is significant is a judgement for the materiality assessment and the assurance provider.

Expect a small number. In SOMA's own example, a hundred-person company running AI hard all year is at a few tens of kilogrammes of CO_2e, not tonnes. The case for the work is rarely the magnitude. It is that someone may ask, and a small number with a method behind it is a better answer than none.

2. Four data tiers, and the evidence to request

Assign the tier per service, from the data you could actually get, never to the company as a whole. Most inventories end up with three tiers in the same table.

Tier	Data in hand	Method	Uncertainty
3	A figure the provider certified for the account	Use it directly	Provider's own. No major provider offers one yet
2a	Exact billed tokens per model	Tokens in millions × factor for class and region	Class range ±48% (A), ±39% (B), ±67% (C), at the world-average grid
2b	Message counts, or seats and an activity rate	Messages × 1,210 tokens, output uplift 2.20 for reasoning, then as 2a	Wider than 2a, with an activity band on estimated tokens
1	An invoice total and nothing else	Spend in EUR × 0.1181 kg CO_2e (EXIOBASE 3.8.2)	±300%

Build the list of services from three places, which are the IT asset register, the finance system's vendor list and the single sign-on log. The services that go missing are the ones on a team credit card. Then ask each provider for tokens by model and by month, for the reporting period.

Tool	What to request	Usual tier
ChatGPT Enterprise or Team	Admin dashboard export of seats and message counts	2b
OpenAI API	Platform console usage export, input and output tokens per model per day	2a
Microsoft 365 Copilot, GitHub Copilot	Admin centre report of assigned seats and active users. No token export	2b
Claude for Work	Seats and activity report	2b
Anthropic API	Console usage export, tokens per model	2a
Gemini for Workspace	Seat count and any activity report the admin console gives	2b, or 1 with invoice only
Azure OpenAI	Azure Monitor token metrics per deployment, plus the invoice. Region is stated	2a
AWS Bedrock	CloudWatch token metrics or invocation logs. The AWS region is the serving region	2a
Google Vertex AI	Cloud Monitoring token counts and billing export. Endpoint location is the serving region	2a
SaaS with embedded AI	The invoice, plus a dated request to the vendor for usage data	1

Keep the export file itself, not a figure copied out of it. If a vendor cannot produce usage, record that you asked and when. That record explains why the line falls back to spend.

The Tier 2b conversion of 1,210 tokens per exchange (928 input, 282 output) is a general business-use assumption. Raise it for coding and long-document work, lower it for short question-and-answer use, and disclose it. The 2.20 output uplift for reasoning modes rests on an author-assumed multiplier, because no public source gives one. Where a tool is reasoning-heavy, get billing tokens and use Tier 2a, which already includes those tokens.

3. The method in five steps

> **kg CO_2e = tokens (millions) × factor (kg CO_2e per million tokens, for the model class and the serving region)**

Step 1. List the services and assign a tier to each. Record vendor, contract type, annual spend, owner and reporting period. That table is the skeleton of the inventory.

Step 2. Turn activity data into tokens. Tier 2a uses billed tokens as exported. Tier 2b converts messages at 1,210 tokens each, with the reasoning uplift on output. Tier 1 skips this step and multiplies spend.

Step 3. Choose the model class. Class A, small (gpt-4o-mini, gpt-4.1-mini, gpt-4.1-nano). Class B, mid (gpt-4o, gpt-4.1). Class C, frontier (gpt-4, gpt-4-turbo, gpt-5). Getting the class wrong is the largest single error in the method, at about four to five times. "Mini" does not always mean small, and a mixture-of-experts model is classed on its active parameters. If an export lists several models, split the tokens and apply each class. SOMA treats an unrecognised model as class C and flags it.

Step 4. Choose the serving region, location-based. Use the region where the provider runs the model, not the client's office. Azure, Bedrock and Vertex name it in the export. For consumer-style services, use the provider's usual serving region and record the assumption, or the global average if even that is unknown. Factors are grid averages for the region, not the provider's renewable contracts. A selection, in kg CO_2e per million tokens:

Region	Class A	Class B	Class C
EU North - Sweden	0.002	0.007	0.029
EU West - France	0.002	0.008	0.033
EU South - Spain	0.005	0.016	0.063
EU-27 average	0.006	0.021	0.081
EU West - Ireland	0.007	0.025	0.093
EU Central - Germany	0.009	0.030	0.115
US East - N. Virginia	0.008	0.026	0.099
Global average	0.012	0.039	0.147
Asia Pacific - Singapore	0.012	0.040	0.151

The table values count electricity, the share of server hardware, data-centre construction and amortised training, at the dataset's default input and output mix. The inference terms are computed per effective token, so a very different mix (for example after the reasoning uplift) gives a somewhat different figure. The methodology page has the term-by-term values, and the full table covers 18 regions.

Step 5. Report the figure with its uncertainty and its sources. Put the band next to the figure, in the same row. Keep three things apart.

- The class range (±48%, ±39%, ±67% at the world-average grid), which is the spread of benchmarked models around the class mean.
- The activity band on estimated tokens at Tier 2b. In SOMA's three-service example it is −50% to +130%.
- A conservative high case, reported beside the central value. It adds about 63.9% (class A) and 63.7% (classes B and C) to the inference footprint.

Then document enough for someone else to rebuild the number without asking you. That means activity data and period, its source file and download date, tier and why, the factor value with its class and region, the factor version (soma-ai-ef-2026.10-dataset-v3.0.0) and DOI, the uncertainty and its assumptions, and whether the line is metered, converted or estimated. Pin one dataset version per reporting year, report gross before offsets, and state the boundary. The factors exclude provider offices and research, network transport and the client's own devices. Embedding, image, speech and video models are outside this factor version and should be reported separately. A development line beyond final training runs is estimated separately and never added to the total.

4. Worked example

Take a company of about 200 people where 72 use AI tools at work. This is the example organisation in SOMA's free calculator ("See an example"), which counts only those 72 users. Staff without AI tools add nothing. A SaaS tool with embedded AI used by the rest would be a separate Tier 1 line.

Team	People	Tools
Engineering	14	GitHub Copilot, Cursor, Claude
Marketing	10	ChatGPT, Claude for documents
Sales	30	Microsoft 365 Copilot
Customer support	18	ChatGPT

The calculator runs at Tier 2b, from published defaults, with a data completeness score of 30 of 100. It states its assumptions. No provider discloses a serving region, so it uses the global-average grid (483 g CO_2e/kWh). Anthropic's apps and API have no EU processing commitment, so global average again. Microsoft names no model for Microsoft 365 Copilot, so it is priced at the frontier class.

Result (calculator, central and range)	Per month	Per year
CO_2e	1.81 kg (907 g to 4.17 kg)	21.8 kg (10.9 to 50.1 kg)
Electricity, including cooling	3.29 kWh	39.4 kWh
Water, reported beside carbon	6.67 L	80.1 L
CO_2e per AI user	25 g	302 g
Development line, separate, not in total	2.15 kg (995 g to 3.65 kg)	

By tool, ChatGPT is 36% of the total, Cursor 24%, Claude 18%, GitHub Copilot 11% and Microsoft 365 Copilot 10%.

How to read it. The annual line is about 22 kg CO_2e, with a range from about half to about 2.3 times the central value. That is a documentation task, not a reduction priority. The figure rests on defaults, so the upgrade path is concrete. Ask for the Microsoft 365 admin report and the ChatGPT message counts to replace the activity defaults, get token exports for Cursor and Claude to move the engineering lines to Tier 2a, and ask each provider where it serves the account. Region alone moves the result a lot. In SOMA's CSRD guide, the same usage comes to 107.1 kg CO_2e from US East - N. Virginia and 26.6 kg from EU North - Sweden.

5. Common mistakes

- **Treating a spend-based figure as the AI figure.** Spend × 0.1181 kg CO_2e per EUR counts money, not electricity. In the method paper's comparison it comes out 1.7 to 12.3 times a usage-based estimate on deflated spend (2.0 to 15.1 times before deflation). Quote that range, not larger ones. Treat a Tier 1 line as a screening figure for the whole subscription, not an AI figure, and ask the vendor for usage next year.
- **Mixing location-based and market-based.** SAEF factors are location-based. A provider's market-based claim, netted for renewable contracts, is a different number. Do not mix the two in one line, and do not call the difference a saving.
- **Ignoring region.** For a class B model, the carbon from electricity per token is more than 13 times higher in Asia Pacific - Singapore than in EU North - Sweden, about 5 times once the whole factor is counted. Record the region and whether it was assumed.
- **Double counting.** AI bought through the client's own cloud account (Azure OpenAI, Bedrock, Vertex) may already sit in a cloud line or a cloud provider's emissions dashboard. Replace that line only where the AI part can be isolated, and remove spend already counted by a token-based line.
- **Wrong model class.** It moves the figure by about four to five times, more than any factor's uncertainty band.
- **Pricing non-chat work at a chat factor.** Embeddings, images, speech and video are out of scope for this factor version.
- **No version on the line.** If the factor source changes silently, last year's figure cannot be rebuilt.

6. Template table

Copy one row per service. The first three rows are SOMA's hypothetical three-service inventory from its GHG Protocol data-tiers guide.

Tool	Users	Tier	Activity data (period, source)	Factor (class)	Region	kg CO_2e	Range	Source / version
API on a cloud platform	n/a	2a	240 M billed tokens, 1 year, usage export	SAEF, class B	EU West - Ireland	5.89	±39% class range	SAEF soma-ai-ef-2026.10-dataset-v3.0.0
Enterprise chat seats	60	2b	60 × 30 × 12 = 21,600 exchanges, reasoning uplift on output	SAEF, class B	EU-27 average	0.97	0.49 to 2.24	SAEF soma-ai-ef-2026.10-dataset-v3.0.0
SaaS with embedded AI	n/a	1	€12,000 invoice, 1 year	0.1181 kg/EUR	n/a	1,417	±300%	EXIOBASE 3.8.2

In that example about 99.5% of the AI footprint sits in the line with the least information behind it. The obvious next step is a usage request to that vendor.

A safe way to describe the basis in a public report is "Prepared under the Scope 3 Standard (2011), with data types tracked against the working draft of March 2026". Do not claim conformance with the revised GHG Protocol standard, which does not exist yet.

7. Sources, citation and limits

Dataset. AI Inference Emission and Resource Factors for Corporate GHG Inventories (SOMA AI Emission Factors, SAEF). Llopis-Garcia, G., SOMA AI, 19 October 2026. CC BY 4.0, reusable with attribution in client deliverables. Concept DOI 10.5281/zenodo.20443585, which resolves to the latest version (this version 10.5281/zenodo.23371447).

Method. Accounting for Purchased AI Inference in Corporate GHG Inventories: A Token-Based Method. Llopis-Garcia, G. (2026), Sustainable Production and Consumption 70, 320-347, doi 10.1016/j.spc.2026.11.003. Free preprint arXiv:2606.10660.

Spend factor. EXIOBASE 3.8.2, Computer and related services, 0.1181 kg CO_2e per EUR, CC BY-SA 4.0, DOI 10.5281/zenodo.5589597.

Pages this brief draws on.

- somaai.earth/ai-footprint/guides/ai-usage-csrd-scope-3
- somaai.earth/ai-footprint/guides/ghg-protocol-data-tiers
- somaai.earth/ai-footprint/guides/why-ai-footprint-estimates-differ
- somaai.earth/ai-footprint/answers/ai-carbon-footprint-by-country
- somaai.earth/ai-footprint/methodology
- somaai.earth/ai-footprint/factors (CSV download)

Free calculator. somaai.earth/ai-footprint/calculator. No sign-up, about two minutes, Tier 2b first estimate.

Honest limits. SAEF factors are class means from public benchmarks of open-weight models, combined with grid averages. Providers do not publish model sizes, so the class is inferred from published tiers and benchmark behaviour. SAEF publishes its boundary, exclusions and uncertainty. It does not certify a cut-off percentage and has no third-party verification. Say so in the method note and let the assurance provider weigh it.

Contact. lili@somaai.earth, somaai.earth. SOMA turns providers' usage exports into a finished Scope 3 Category 1 entry, with tier, factor, version, sources and uncertainty attached.