Methodology and Dataset Documentation of the Industrial Policy Monitor

Kiel Institute for the World Economy

Version 1.1 - updated 05/12/2024

Table of Contents

1 Overview and scope 1
2 Source corpus 1
3 Retrieval 2
4 Relevance screening 2
5 Extraction and coding 3
6 Consolidation into measures 4
7 Sample definition and counting rules 5
8 Quality assurance and validation 6
9 Limitations 7
References 7
Annex 8

1 Overview and scope

The Industrial Policy Monitor, compiled by the Industrial Policy Lab at the Kiel Institute for the World Economy, is a media-based dataset of government industrial policy measures. It converts national press coverage of six sectors in eleven countries into a structured, English-language inventory of concrete government actions. The resulting data measure reported policy activity and public debate; they are not a legal registry or a census of the policies governments adopt. This appendix documents data version 2e4feecdeda8, which covers press reports up to 11 November 2026 and whose records were last changed on 16 November 2026.

The country sample comprises Germany, France, Italy, Spain, the Netherlands, Poland, Czechia, Estonia, the United States, the United Kingdom and South Korea. It thus combines the four largest EU economies, four smaller member states that differ in size and region, and three advanced economies outside the EU. The six sectors are electric vehicles and automotive, steel and metals, chemicals, semiconductors, hydrogen, and critical minerals. Together they cover the energy transition of heavy industry and transport, technology and economic security, and the raw-material inputs on which these depend, and they span the full range of instruments, from purchase premiums and contracts for difference to tariffs, export controls and stockpiles.

For all 66 country–sector pairs, the press is read from 1 January 2026 to 11 November 2026. For ten pairs, the archive is additionally read from 1 January 2024: electric vehicles in Germany, France, the United Kingdom and the United States; steel in Germany, the United Kingdom and the United States; and semiconductors in Germany, the United States and South Korea.

The unit of analysis is the *measure*, defined as one concrete government action: one instrument for one group of beneficiaries, as the government decides it. Examples are a purchase premium for cars, a tariff, or an R&D grant scheme. A plan or act that bundles several instruments is treated as a package; its constituent measures are counted, but the package itself is not. Measures may belong to an *initiative*, that is, a named programme or act such as an EU regulation or a national strategy. A measure is assigned to an initiative only when the press names it as part of one, and only EU initiatives span countries. The raw input is the *article*, a single press item from one of the listed outlets.

2 Source corpus

All articles are drawn from the LexisNexis press archive, a licensed full-text news archive. The corpus was built through the LexisNexis Web Services API (WSApi), which provides programmatic access to LexisNexis Licensed News and Legal News content.

Each outlet on a country list is assigned to a tier. Tiers T1 to T3 comprise the national general press (dailies, weeklies and broadcasters), national business titles and newswires; together they form the *core press*. Tiers T4 and T5 comprise trade and regional titles.

Each country is read in its domestic language wherever the archive carries the domestic press. The exceptions are South Korea and Czechia, which are read through English-language titles only, and Poland and Estonia, which are read in the domestic language and additionally through English-language services; for Estonia a Russian-language edition is also included. Table A.2 lists every outlet read, by country and tier.

3 Retrieval

For each country–sector pair, a search query retrieves every article in the listed outlets that mentions the sector's products or industry. The queries are built from sector terms rather than policy terms. The only exceptions are a small number of policy names retained in the German and French queries and in the Dutch critical-minerals query. Each query consists of 8 to 21 exact phrases joined by OR; Table A.3, at the end of this appendix, lists them for every country–sector pair. Where a language filter is applied in the archive, it matches the country's reading language; for Poland, Czechia and Estonia no filter is applied, the outlet list sets the language, and the queries combine domestic-language and English terms. Policy relevance is therefore not determined by the search string but by the screening step described in Section A.4. This design avoids conditioning the sample on a predefined list of instruments and allows the pipeline to identify instruments that were not anticipated.

Articles are retrieved from the archive week by week, and the corpus is extended at each periodic update. An article returned by two sector queries is counted once in the retrieved total. Within each country–sector corpus, copies of the same headline published on the same day are removed before screening (50,179 in all). A wire story republished under a new headline or on a later day is treated as a separate article. In total, 337,486 articles were retrieved, including the 2024–25 reading of the ten pairs read from 2024.

4 Relevance screening

Each retrieved article is read in its original language by an AI model (DeepSeek V4 Flash and, from October 2026, DeepSeek V4.1 Flash, accessed through the OpenRouter API; see Section A.8 for the versions)[1], which makes a binary decision to retain or discard it. An article is retained if a specific, current government action aimed at the sector is its main subject. The model is given a definition of industrial policy adapted from Juhász, Lane and Rodrik (2023), according to which industrial policies are government policies that "explicitly target transforming the structure of economic activity in pursuit of a public goal".

The action must be taken by the tracked country, at the national, regional or local level, or in an agreement to which that country is party, or, for EU member states, by the European Union.

[1] deepseek-v4-flash-0731 (5 August to 10 October 2026), deepseek-v4.1-flash (11 to 18 October 2026) and deepseek-v4.1-flash-1017 (from 19 October 2026). Articles dated 2026 were screened with the first version if retrieved up to late September and with the third thereafter; the 2024–25 history of the ten pairs was screened with the second (11 to 13 October 2026).

Actions of other governments are excluded. A plant or project located abroad is excluded even when the tracked government finances it, for example through a public loan for a mine in another country; this rule is enforced in the review stage. Of the 337,486 retrieved articles, 16,060 articles (4.8%) pass the relevance screen; 12,684 (3.8%) support at least one measure.

5 Extraction and coding

The same AI model, run with extended reasoning, extracts every concrete measure reported in each retained article. A single written definition of a measure, namely one instrument for one group of beneficiaries, governs both extraction and the subsequent merging of records. The definitions given to the model in the screening, extraction and merging prompts are listed in Table A.4. Because each article is processed separately, a widely reported action is extracted many times in different wordings; these records are consolidated in the step described in Section A.6.

For each measure, the model records the instrument, the amount where one is stated, the deciding body, the implementing body, the beneficiaries, the stage of the measure, any conditions attached to it, and an evidence quote from the article. All content fields are written in English, while evidence quotes are retained in the language of the original article. Each measure is then coded on the following dimensions.

Sector. Each record is assigned to exactly one of the six sectors according to fixed boundary rules. Battery cell plants are coded as electric vehicles rather than critical minerals; ammonia and fertiliser plants as chemicals rather than hydrogen; and direct-reduction steelmaking as steel rather than hydrogen. Critical minerals cover only the upstream stages of mining, refining, processing, stockpiling and recycling, together with raw-materials agreements. A rule that applies to several sectors, such as a carbon border charge or an industrial electricity price, is recorded once under each sector to which it applies.

Policy goal. Each measure is assigned the public objective as reported in the press, drawn from ten categories: green transformation, industrial competitiveness, resilience and security, defence and space, employment and labour, regional development, digital transformation, physical infrastructure, trade agreements, and other public goals. Most measures carry a single goal and a few carry two or three. Defence and space includes military procurement, support for the defence industry and dual-use technology; trade agreements include export promotion and market access.

Status. Status records the most recent stage reported by the press. A measure is *in discussion* when it is debated, proposed or negotiated without formal commitment; *announced* when it is formally committed or decided but not yet implemented; *active* when implementation has begun, that is, when funds are flowing, construction is under way or rules are in force; *cancelled* when it was previously announced or active and has since been withdrawn or abandoned; and *expired* when it was previously active and has ended or lapsed. Where the reporting does not

identify a stage, status is coded as not stated. An adopted law with a fixed start date is coded as active from that date unless a delay is reported. Status is not verified against official gazettes.

Level of government. This variable records which body decided or announced the measure, not which body implements it, and takes the values EU, national, regional, or not stated. Named ministries and agencies are coded as national, and municipalities are grouped with the regions. Where the press names them, the deciding and implementing bodies are recorded by name.

Value. Value records the public money behind a measure as reported in the press. It includes grants, tax reliefs, public loans, guarantees and equity, including those provided by state and policy banks, that a government commits or spends, or proposes while the measure is in discussion. A reported figure is excluded when it does not constitute such support. This applies to a company's own investment or pledge; to cuts and savings; to taxes, levies and penalties collected by the state, trade volumes and agencies' running budgets; to lending ceilings, undrawn guarantees and conditional loans that never closed; to the total size of a public lending fund, whose individual awards are counted instead; to schemes financed by levies, consumers or operators; and to party or campaign proposals and members' bills that have not passed.

Several further rules prevent double counting. A programme is counted once, so that where its total is recorded its individual awards add nothing. An EU-wide or multi-country amount is counted only for the country's own share. For a cancelled programme, only the amount paid out is counted. For an economy-wide package, only a stated sector share is counted, except that programmes designed for energy-intensive industry retain their full amount in steel and chemicals. A figure printed at different scales in different outlets is treated as one figure, and where outlets report different figures the record retains one of them rather than a sum. Annual figures are flagged as such, and non-monetary figures such as numbers of vehicles, gigawatts or jobs are stored separately in their own units. Excluded figures remain in the measure's description but do not enter the value field. Value is missing for most measures.

Dates. The date of first report is the date of the measure's earliest article in the data. The status history records the dates the press gives for each stage, often at month or year precision only.

6 Consolidation into measures

First, retained articles are grouped by the similarity of a short description of each article, consisting of its headline and the screening model's one-sentence rationale, policy name, list of measures and policy goal. The descriptions are embedded with a multilingual sentence-embedding model (paraphrase-multilingual-MiniLM-L12-v2, 384 dimensions); an article joins the existing group whose centroid it is closest to if the cosine similarity is at least 0.85, and otherwise opens a new group. The measures found in the articles of one group are compared and merged month by month. Because a widely reported programme may still fall into several groups, a second step merges, within each country and sector, all repeated reports of the same measure into a single record on the basis of the model's judgement. One record is chosen to

represent the merged measure, which takes that record's status unless review shows that later reports document a further stage. Each canonical measure stores the number of press-reported records from which it was consolidated; most measures rest on one or two articles.

Third, each periodic update carries the existing records forward, so that new coverage is attached to existing measures and status histories accumulate over time. Fourth, measures are linked to the initiative of which the press reports them to be part, for example a single EU act reported in four languages. Initiative names are canonical English forms chosen by the pipeline. In the present data version, 2,547 of the 5,768 measures in the full dataset (44 per cent) belong to a named initiative.

7 Sample definition and counting rules

Trade and regional titles vary greatly in number across countries and would inflate counts unevenly. Cross-country analysis therefore uses the *curated set*, which comprises every measure with at least one article from a core outlet (tiers T1 to T3) dated on or after 1 January 2026. Measures reported only by trade or regional titles remain in the *full dataset* but are excluded from the curated set. This restriction holds the type of outlet and the observation window constant across countries, although it does not equalise the depth of each country's core-press coverage.

All counts refer to canonical measures, not to articles or packages. An EU act is recorded under each member state whose press reported it, and a cross-sector rule under each sector to which it applies, so that totals summed across countries or sectors may count the same act more than once. Because merging removes duplicate records, the number of measures can fall between data versions without any loss of information. An article that supports several measures is counted once in article totals, while a wire story filed in two country or sector corpora is counted once in each. A country–sector cell without measures indicates that the corpus was read and no measure was found; a failed pipeline stage produces no data and is reported as such rather than recorded as zero.

A measure is assigned to a period if at least one of its articles is dated within it, including when the measure was first reported earlier. Periods are counted back from the date of the newest article, 11 November 2026, rather than from the date of access. A measure is classified as new in a period if its first article in the data falls within that period. The date of first report is a press observation and not the date of decision or enactment. For country–sector pairs read only from 1 January 2026, a measure first observed early in 2026 may have been under discussion well before.

Table A.1 reports the number of observations at each stage of the pipeline by country. It shows how much of the cross-country variation in measures reflects differences in corpus size and outlet depth.

Table A.1. From retrieved articles to curated measures, by country

Country	Core outlets	Articles retrieved	After de-dupli-cation	Screened in	Support-ing a measure	Records before consoli-dation	Measures (full)	Measures (curated)
Germany	13	90,109	79,091	4,740	3,904	2,823	1,138	275
France	15	41,252	38,330	1,456	1,288	1,613	702	263
Italy	11	26,518	19,633	1,534	1,018	1,313	651	554
Spain	9	20,165	17,473	1,002	756	1,016	512	398
Netherlands	10	1,020	1,012	49	39	55	36	36
Poland	4	2,624	2,572	125	89	110	60	60
Czechia	3	672	672	39	23	37	26	26
Estonia	5	1,539	1,077	62	33	53	35	35
United Kingdom	12	55,678	48,236	1,859	1,494	1,179	454	103
United States	18	80,837	62,754	4,532	3,614	3,333	1,726	238
South Korea	6	17,072	16,457	662	426	817	428	279
Total	**106**	**337,486**	**287,307**	**16,060**	**12,684**	**12,349**	**5,768**	**2,267**

Notes: Data version 2e4feecdeda8. Core outlets are core-press outlets with at least one retrieved article. Article counts are distinct per country: an article returned by two sector queries is counted once, although it appears in both cells of the country–sector file. De-duplication removes same-day copies of a headline within a country–sector corpus. Records before consolidation are counted after the merge within article groups and before the merge within country and sector. For Germany, France, the United Kingdom, the United States and South Korea, all columns except curated measures include the 2024–25 reading. In the Netherlands, Poland, Czechia and Estonia nearly all outlets are core press, so all measures are curated. Measure totals count an EU act once per reporting member state and a cross-sector rule once per sector.

8 Quality assurance and validation

Quality control combines a second model pass, rule-based checks and review. The second pass re-examines every merged group and splits off records that do not describe the same action. At each update, rule-based checks flag suspect records, such as dates in the future or company funds coded as public money; flagged records are corrected or confirmed, and all corrections are kept in a file that is reapplied at every update. At the time of writing, this file holds 2,615 field corrections and 91 splits of wrongly merged records, and 1,822 identifiers have been retired through merging or removal. Every measure retains its evidence quote and the list of supporting articles; full texts are not redistributed because the archive is licensed.

No hand-coded benchmark is available; the checks below were run by AI agents independent of the extraction model and by re-running the models. Re-screening a stratified sample of 2,975 stored articles with the current model agreed with the stored decision for 98.7 per cent of articles (re-weighted to the corpus; Cohen's kappa 0.86). Nearly all disagreements were articles the current model no longer retains, and in 40 adjudicated cases the current decision was judged correct in 27 and the stored one in 7. Applied to the whole corpus, re-screening would lower

the number of core-press articles supporting a measure by 13 to 19 per cent but leave ratios between years almost unchanged (2026 relative to 2025: 0.99, 95 per cent interval 0.89 to 1.11). Re-running the same model version changes the decision for about one retained article in ten, all of them borderline cases. Recall on discarded articles is largely unmeasured (2 of 1,500 discards were retained on re-screening). A re-reading of 76 of the 354 initiative links added in review found no errors. The run-to-run stability of extraction, merging and coding has not been assessed.

9 Limitations

The data measure press attention rather than policy volume. Corpus size, language, outlet mix and archive depth differ across countries, and a country read through more outlets will show more measures. Cross-country comparisons should therefore rely on shares within the curated set, for example by status, instrument or policy goal, rather than on absolute counts. Even within the curated set, absolute counts partly reflect the depth of each country's core-press coverage.

Status reflects the last reported stage rather than verified implementation; a measure that remains at the announced stage may be stalled or may simply have ceased to attract coverage. Amounts are claims made in the press rather than verified budget lines, are missing for most measures, and should not be summed. Finally, the number of articles reporting a measure indicates its political salience rather than its size or economic importance.

References

Juhász, R., N. Lane and D. Rodrik (2023). The New Economics of Industrial Policy. NBER Working Paper 31538.

Annex

Table A.2. Outlets read, by country and tier

Tier	Outlets (articles retrieved)
Germany (German)	
T1	Frankfurter Allgemeine Zeitung (9,642); Stern (8,729); Frankfurter Rundschau (3,794); DER SPIEGEL (online) (2,220); Der Tagesspiegel (2,067); taz, die tageszeitung (2,007); FAZ+ (1,833); Frankfurter Allgemeine Sonntagszeitung (959); FOCUS - Das Nachrichtenmagazin (533); Frankfurter Allgemeine Zeitung final version (R2) (319); SPIEGEL+ (311)
T2	FOCUS MONEY (1,201)
T3	dpa-AFX ProFeed (27,626)
T4	Energie & Management (2,176); Elektronik Praxis (1,878); Automobil Industrie (1,820); Automobilwoche (1,174); VDI nachrichten (813)
T5	Rheinische Post (13,633); Sächsische Zeitung (4,035); Kölner Stadt-Anzeiger (3,630); General-Anzeiger Bonn (3,188); Leipziger Volkszeitung (3,077); Nürnberger Nachrichten (1,540); Rheinische Post Online (752); General-Anzeiger Bonn Online (197)
France (French)	
T1	Le Figaro (2,989); Le Monde (2,372); Libération (1,915); LExpress.fr (1,036); La Croix (468); Liberation.fr (430); L'Humanité (407); L'Express (359); FRANCE 24 (French) (359); Le Point (352); Marianne (105)
T2	Le Figaro Economie (2,964); La Tribune Dimanche (2,394); La Tribune.fr (1,511); Challenges (83)
T5	La Voix du Nord (5,008); Le Progrès (4,673); L'Est Républicain (3,768); Le Télégramme (3,679); Midi Libre (2,724); Nice-Matin (2,377); Sud Ouest et Sud Ouest Dimanche (1,834); La Dépêche du Midi (1,752); 20 Minutes (11)
Italy (Italian)	
T1	Corriere della Sera (Italy) (3,056)
T2	ItaliaOggi (899); MF (820); Milano Finanza (381); FTAOnline (353); ItaliaOggi7 (104)
T3	ANSA Notiziario Generale in Italiano (7,549); ANSA Financial News (4,774); Alliance News Italian Service (1,345); ANSA Business News (590); ANSAmed - Italian (77)
T5	Il Resto del Carlino (Italy) (3,446); La Nazione (Italy) (3,256); Il Giorno (Italy) (1,829)
Spain (Spanish)	
T1	El Pais - Spanish (online) (2,369); El Confidencial (946); El Pais (762)
T2	El Economista, Spain (4,491); Expansion (3,329); Expansion Online (1,824); Cinco Días (online) (1,348); Cinco Días (1,091); Actualidad Económica (389)
T4	EFE - Spanish Newswire (6,060); dpa internacional (Spanischer Dienst) (468)
Netherlands (Dutch)	
T1	Reformatorisch Dagblad (263); Nederlands Dagblad (248); EW Magazine (178); De Groene Amsterdammer (79); HP/De Tijd (37); Vrij Nederland (23)
T2	Beleggers Belangen (239); Quotenet (62); Forum (29); Quote (24)
Poland (Polish, English)	
T1	Gazeta Wyborcza (1,835)
T3	PAP Market Insider (608); PAP English News Service (204); Polish News Bulletin (193)
Czechia (English)	
T3	CTK Business News (526); CTK National News Wire (131); Intellinews - Czech Republic This Week (74)
T4	Business Wire Czech (17)
Estonia (Estonian, English, Russian)	
T1	Postimees Estonian (1,072); Postimees Russian (19); Postimees English (13)

T2	Capital (English) (54)
T3	Baltic News Service (BNS) English (494)
United Kingdom (English)	
T1	The Guardian (London) (9,503); The Daily Telegraph (London) (6,324); Scottish Daily Mail (2,677); The Daily Mail (London) (2,448); The Sunday Telegraph (London) (903); The Observer - London (Online) (603)
T2	The Investors Chronicle (946); City A.M. (627)
T3	Alliance News UK (3,650); PA UK & Ireland National Newswire (2,817); PA Newswire: Financial Newswire (754); BBC Monitoring: International Reports (2)
T4	Electronics Weekly (1,141); New Civil Engineer (652); New Scientist (297); Utility Week (195); Chemical Industry Digest (1)
T5	The Western Mail (3,671); Express and Star (2,069); The Herald (Scotland) Online (2,051); Yorkshire Post (1,901); The Journal (Newcastle, UK) (1,744); Daily Record and Sunday Mail (1,595); Derby Telegraph (1,404); Hull Daily Mail (1,369); Coventry Evening Telegraph (1,140); Scotsman (1,137); Manchester Evening News (1,102); Grimsby Telegraph (1,079); The Herald (Scotland) Print (1,011); Liverpool Echo (947); Yorkshire Evening Post (894); Evening Times (Glasgow) (536); Scunthorpe Telegraph (501); Birmingham Post (455); PA Newswire: Scotland (265); PA Regional Newswire of English Regions: YORKS AND HUMBER (205); The Sunday Herald (Glasgow) (133); PA Regional Newswire of English Regions: NORTH EAST (49)
United States (English)	
T1	USA Today Online (6,326); The New York Times - International Edition (3,057); Los Angeles Times (2,886); USA Today (1,904); Los Angeles Times Online (1,842)
T2	Business Insider (4,084); Forbes.com (Forbes, Incorporated) (3,019); Axios (2,794); Politico.com (196); National Journal (188); The Hill (184); CQ Budget Tracker (116); Forbes (101); Congressional Quarterly Magazine (48)
T3	The Associated Press (8,573); Associated Press International (6,916); Associated Press Financial Wire (5,217); Associated Press State & Local (2,606)
T4	Inside U.S. Trade's Daily Report (1,861); Inside U.S. Trade (1,386); Chemical Engineering (442); Greenwire (403); Industry Week (327); Energywire (300); ClimateWire (280); E&E News PM (98); Government Executive (18)
T5	The Detroit News (Michigan) (3,544); Detroit Free Press Online (Michigan) (3,397); Pittsburgh Post-Gazette (3,249); Chicago Tribune (3,061); The Boston Globe (3,046); Detroit Free Press (Michigan) (2,577); The San Francisco Chronicle (California) (1,291); The Palm Beach Post (1,283); The Dallas Morning News (1,282); The Columbus Dispatch (Ohio) (1,264); The Times Union (Albany, NY) (1,237); The Tennessean (Nashville, Tennessee) (1,206); The Houston Chronicle (1,181); The Atlanta Journal-Constitution (1,162); The Philadelphia Inquirer (1,017); The Denver Post: Web Edition Articles (Colorado) (853); The Post-Standard (Syracuse, NY) (795); The Indianapolis Star (Indiana) (702); The Plain Dealer (617); Crain's Detroit Business (431); The State (Columbia, South Carolina) (335); Crain's Chicago Business (258)
South Korea (English)	
T1	The Korea Herald (4,367); Korea Times (2,957)
T2	The Korea Economic Daily Global Edition- English (1,777)
T3	Korea News Gazette (5,119); South Korea Daily Report (4,606); Intellinews - South Korea Today (141)

Notes: Outlet names as in the LexisNexis archive. Reading languages in parentheses. Tiers: T1 national general press; T2 national business press; T3 newswires; T4 trade, specialist and foreign-language wire services; T5 regional press. Tiers T1–T3 form the core press. Figures in parentheses are articles retrieved summed over the six sector corpora, so an article returned by two sector queries is counted twice. Süddeutsche Zeitung, Handelsblatt, Les Echos and the AFP wire are not on the lists.

Table A.3. Search terms by country and sector

Sector (terms)	Search terms
Germany (archive language filter: German)	
EVs & automotive (18)	Elektroauto; Elektromobilität; E-Auto; Ladeinfrastruktur; Ladesäule; Batteriefabrik; Batterieproduktion; electric vehicle; charging station; gigafactory; E-Fuels; Autozölle; Batteriezellfertigung; Kaufprämie; Autoindustrie; Automobilindustrie; Autohersteller; Autobranche

Steel & metals (16)	Stahl; Stahlproduktion; Stahlindustrie; Stahlwerk; grüner Stahl; Hochofen; Direktreduktion; Stahlimport; Stahlzoll; klimaneutraler Stahl; CBAM; CO2-Grenzausgleich; Grenzausgleichsmechanismus; Stahlbranche; Stahlkonzern; Stahlhersteller
Chemicals (14)	Chemieindustrie; Chemiesektor; Grundstoffchemie; Spezialchemie; Petrochemie; Kunststoffindustrie; Düngemittel; Chemiepark; chemisches Recycling; Industriegase; Industriestrompreis; Chemie; Chemiebranche; Chemiekonzern
Semi-conductors (14)	Halbleiter; Chipindustrie; Mikroelektronik; Chipfabrik; Halbleiterwerk; Wafer; Foundry; Chipproduktion; Halbleiterförderung; EU Chips Act; Exportkontrollen; Halbleiterbranche; Chiphersteller; Chipkonzern
Hydrogen (14)	Wasserstoff; grüner Wasserstoff; Elektrolyseur; Elektrolyseure; Brennstoffzelle; Brennstoffzellen; Wasserstoffnetz; Kernnetz; Wasserstofftankstelle; Ammoniak; H2Global; IPCEI Wasserstoff; Wasserstoffwirtschaft; Wasserstoffbranche
Critical minerals (15)	Seltene Erden; kritische Rohstoffe; Lithium; Kobalt; Graphit; Gallium; Germanium; Permanentmagnete; Rohstofffonds; Rohstoffabkommen; Lithiumabbau; Raffineriekapazität; Rohstoffbranche; Bergbaukonzern; Bergbau

France (archive language filter: French)

EVs & automotive (16)	voiture électrique; véhicule électrique; électromobilité; borne de recharge; infrastructure de recharge; batterie; usine de batteries; gigafactory; bonus écologique; prime à la conversion; leasing social; e-carburants; droits de douane automobiles; industrie automobile; constructeurs automobiles; secteur automobile
Steel & metals (13)	acier; sidérurgie; production d'acier; aciérie; acier vert; haut fourneau; réduction directe; importations d'acier; droits de douane acier; CBAM; MACF; sidérurgiste; filière acier
Chemicals (12)	industrie chimique; chimie; chimie de base; chimie de spécialité; pétrochimie; plasturgie; engrais; site chimique; recyclage chimique; gaz industriels; secteur chimique; chimiste
Semi-conductors (13)	semi-conducteurs; puces électroniques; microélectronique; usine de puces; fonderie; wafer; fabrication de puces; industrie des semi-conducteurs; Chips Act; composants électroniques; contrôles à l'exportation; fabricants de puces; filière semi-conducteurs
Hydrogen (11)	hydrogène; hydrogène vert; électrolyseur; électrolyseurs; pile à combustible; piles à combustible; station hydrogène; ammoniac; hydrogène décarboné; filière hydrogène; secteur de l'hydrogène
Critical minerals (13)	terres rares; matières premières critiques; lithium; cobalt; graphite; gallium; germanium; aimants permanents; minerais critiques; métaux critiques; raffinage de lithium; secteur minier; industrie minière

Italy (archive language filter: Italian)

EVs & automotive (17)	auto elettrica; auto elettriche; veicolo elettrico; veicoli elettrici; mobilità elettrica; colonnina di ricarica; colonnine di ricarica; stazione di ricarica; punti di ricarica; infrastruttura di ricarica; gigafactory; fabbrica di batterie; celle per batterie; industria automobilistica; settore automotive; case automobilistiche; filiera automotive
Steel & metals (15)	acciaio; acciaieria; acciaierie; siderurgia; altoforno; altiforni; forno elettrico; riduzione diretta; preridotto; minerale di ferro; rottami ferrosi; laminatoio; industria siderurgica; settore siderurgico; filiera siderurgica
Chemicals (16)	chimica di base; chimica fine; petrolchimica; impianto chimico; impianti chimici; polo chimico; cracking; fertilizzanti; concimi; gas industriali; riciclo chimico; materie plastiche; plastica; industria chimica; settore chimico; filiera chimica
Semi-conductors (15)	semiconduttore; semiconduttori; microchip; microelettronica; circuito integrato; circuiti integrati; fabbrica di chip; fonderia di chip; wafer; litografia; packaging avanzato; componenti elettronici; produttori di chip; industria dei chip; filiera dei chip
Hydrogen (12)	idrogeno; idrogenodotto; idrogenodotti; elettrolizzatore; elettrolizzatori; elettrolisi; cella a combustibile; celle a combustibile; ammoniaca verde; Hydrogen Valley; filiera dell'idrogeno; economia dell'idrogeno
Critical minerals (15)	materie prime critiche; minerali critici; terre rare; litio; cobalto; grafite; gallio; germanio; neodimio; magneti permanenti; estrazione mineraria; raffinazione; industria mineraria; settore minerario; filiera mineraria

Spain (archive language filter: Spanish)

EVs & automotive (21)	coche eléctrico; coches eléctricos; vehículo eléctrico; vehículos eléctricos; movilidad eléctrica; punto de recarga; puntos de recarga; infraestructura de recarga; electrolinera; electrolineras; gigafactoría; fábrica de baterías; planta de baterías; celdas de batería; celdas de baterías; producción de baterías; industria automovilística; sector del automóvil; sector de la automoción; fabricantes de automóviles; industria del motor
Steel & metals (20)	acero; aceros especiales; acería; acerías; siderurgia; siderúrgica; siderúrgicas; alto horno; altos hornos; horno alto; horno eléctrico; horno de arco eléctrico; reducción directa; mineral de hierro; chatarra férrica; laminación; industria siderúrgica; sector siderúrgico; sector del acero; industria del acero

Chemicals (18)	química básica; química fina; petroquímica; polo químico; polo petroquímico; complejo petroquímico; planta química; plantas químicas; fertilizantes; gases industriales; reciclaje químico; reciclado químico; plásticos; industria química; sector químico; empresas químicas; industria del plástico; sector petroquímico
Semi-conductors (19)	semiconductor; semiconductores; microchip; microchips; microelectrónica; circuito integrado; circuitos integrados; fábrica de chips; fábricas de chips; planta de chips; obleas de silicio; litografía; diseño de chips; encapsulado avanzado; componentes electrónicos; industria de los semiconductores; sector de los semiconductores; fabricantes de chips; industria de los chips
Hydrogen (18)	hidrógeno; electrolizador; electrolizadores; electrólisis; electrolisis; pila de combustible; pilas de combustible; amoniaco verde; amoníaco verde; hidrogenoducto; hidrogenoductos; hidroducto; hidroductos; hidrogenera; hidrogeneras; economía del hidrógeno; sector del hidrógeno; cadena de valor del hidrógeno
Critical minerals (18)	materias primas críticas; materias primas fundamentales; minerales críticos; minerales estratégicos; tierras raras; litio; cobalto; grafito; galio; germanio; neodimio; imanes permanentes; extracción minera; minería metálica; industria minera; sector minero; industria extractiva; empresas mineras

Netherlands (archive language filter: Dutch)

EVs & automotive (12)	elektrische auto; elektrische auto's; elektrisch rijden; laadpaal; laadpalen; laadinfrastructuur; batterijfabriek; accufabriek; elektrisch vervoer; auto-industrie; autofabrikant; autofabrikanten
Steel & metals (9)	staalindustrie; staalfabriek; hoogoven; hoogovens; staalproductie; groen staal; staalproducent; staalsector; staalbedrijf
Chemicals (10)	chemische industrie; chemiebedrijf; petrochemie; kunstmest; kunstmestfabriek; chemiecluster; naftakraker; basischemie; chemiesector; chemieconcern
Semi-conductors (14)	halfgeleider; halfgeleiders; chipfabriek; chipindustrie; chipmachine; chipmachines; microchip; microchips; wafer; wafers; lithografie; chipsector; chipbedrijf; chipbedrijven
Hydrogen (9)	waterstof; groene waterstof; waterstoffabriek; elektrolyser; elektrolysers; waterstofnetwerk; waterstofeconomie; waterstofsector; waterstofprojecten
Critical minerals (10)	kritieke grondstoffen; kritieke materialen; zeldzame aardmetalen; lithium; kobalt; grafiet; grondstoffenstrategie; permanente magneten; mijnbouw; mijnbouwsector

Poland (no language filter; Polish and English terms)

EVs & automotive (21)	samochody elektryczne; samochodów elektrycznych; auta elektryczne; pojazdów elektrycznych; elektromobilność; elektromobilności; stacje ładowania; stacji ładowania; fabryka baterii; produkcja baterii; electric vehicle; electric vehicles; electric car; charging station; electromobility; gigafactory; przemysł motoryzacyjny; przemysłu motoryzacyjnego; branża motoryzacyjna; branży motoryzacyjnej; automotive industry
Steel & metals (21)	stal; produkcja stali; hutnictwo; hutnictwa; huta stali; wielki piec; wielkiego pieca; stalownia; walcownia; piec łukowy; steel; steel mill; steelworks; steel production; blast furnace; electric arc furnace; przemysł stalowy; przemysłu stalowego; przemysł hutniczy; sektor stalowy; steel industry
Chemicals (21)	petrochemia; petrochemii; nawozy; nawozów; tworzywa sztuczne; tworzyw sztucznych; zakłady chemiczne; zakładów chemicznych; chemikalia; kompleks petrochemiczny; chemical plant; petrochemical; petrochemicals; fertilizers; fertilizer; plastics; przemysł chemiczny; przemysłu chemicznego; branża chemiczna; branży chemicznej; chemical industry
Semi-conductors (21)	półprzewodniki; półprzewodników; chipy; chipów; mikroelektronika; mikroelektroniki; układy scalone; układów scalonych; mikroprocesory; mikroprocesorów; semiconductor; semiconductors; microchip; microchips; chip plant; chip factory; przemysł półprzewodnikowy; przemysłu półprzewodnikowego; producenci chipów; semiconductor industry; chipmakers
Hydrogen (21)	wodór; wodoru; elektrolizery; elektrolizerów; ogniwa paliwowe; ogniw paliwowych; zielony amoniak; zielonego amoniaku; dolina wodorowa; autobusy wodorowe; hydrogen; electrolyzer; electrolysers; fuel cell; fuel cells; green ammonia; gospodarka wodorowa; gospodarki wodorowej; technologie wodorowe; hydrogen economy; hydrogen industry
Critical minerals (21)	surowce krytyczne; surowców krytycznych; ziemie rzadkie; ziem rzadkich; litu; kobalt; kobaltu; miedź; miedzi; grafit; critical raw materials; critical minerals; rare earths; rare earth; lithium; copper; przemysł wydobywczy; przemysłu wydobywczego; sektor wydobywczy; mining industry; mining sector

Czechia (no language filter; Czech and English terms)

EVs & automotive (21)	elektromobily; elektromobilů; elektromobilita; elektromobility; elektrická auta; elektrických aut; dobíjecí stanice; dobíjecích stanic; nabíjecí stanice; továrna na baterie; výroba baterií; electric vehicle; electric vehicles; electric cars; charging stations; gigafactory; automobilový průmysl; automobilového průmyslu; autoprůmysl; automobilky; automotive industry
Steel & metals (21)	ocel; oceli; ocelářství; huť; hutě; železárny; železáren; ocelárny; vysoká pec; vysoké pece; obloukové pece; steel; steelmaker; steelworks; steel mill; blast furnace; hutní průmysl; hutního průmyslu; hutnictví; steel industry; steel sector

Chemicals (20)	chemička; chemičky; petrochemie; hnojiva; hnojiv; plasty; plastů; chemická výroba; chemické výroby; etylenová jednotka; chemical plant; petrochemicals; petrochemical; fertilizer; fertiliser; plastics; chemický průmysl; chemického průmyslu; chemickém průmyslu; chemical industry
Semi-conductors (20)	čipy; čipů; polovodiče; polovodičů; mikročipy; mikročipů; mikroelektronika; mikroelektroniky; karbidu křemíku; semiconductor; semiconductors; chipmaker; chip plant; microchips; silicon carbide; polovodičový průmysl; polovodičového průmyslu; výrobci čipů; semiconductor industry; chip industry
Hydrogen (20)	vodík; vodíku; vodíkem; elektrolyzér; elektrolyzéry; elektrolýza; palivové články; palivových článků; vodíkové údolí; zelený amoniak; hydrogen; electrolyser; electrolyzer; fuel cells; green ammonia; Hydrogen Valley; vodíková ekonomika; vodíkové technologie; vodíkových technologií; hydrogen economy
Critical minerals (20)	kritické suroviny; kritických surovin; vzácné zeminy; vzácných zemin; nerostných surovin; lithium; lithia; kobalt; kobaltu; grafit; permanentní magnety; critical raw materials; critical minerals; rare earths; rare earth; cobalt; těžební průmysl; těžebního průmyslu; mining industry; mining sector
Estonia (no language filter; Estonian and English terms)	
EVs & automotive (11)	elektriauto; elektriautod; laadimispunkt; laadimistaristu; akutehas; electric vehicle; electric vehicles; charging station; battery plant; autotööstus; car industry
Steel & metals (8)	terasetööstus; terasetehas; metallitööstus; steel industry; steel plant; steelmaker; terasesektor; steel sector
Chemicals (11)	keemiatööstus; keemiatehas; väetis; põlevkivi; põlevkivitööstus; chemical industry; chemical plant; fertiliser; oil shale; keemiasektor; chemicals sector
Semi-conductors (11)	pooljuht; pooljuhid; kiibitehas; mikrokiip; elektroonikatööstus; semiconductor; semiconductors; microchip; chip plant; kiibitööstus; chip industry
Hydrogen (9)	vesinik; roheline vesinik; vesinikutehas; elektrolüüser; hydrogen; green hydrogen; electrolyser; vesinikusektor; hydrogen sector
Critical minerals (11)	kriitilised toormaterjalid; haruldased muldmetallid; fosforiit; maavarad; critical minerals; critical raw materials; rare earth; rare earths; phosphorite; kaevandus; mining industry
United Kingdom, United States and South Korea (identical queries; archive language filter: English)	
EVs & automotive (18)	electric vehicle; electric vehicles; electric car; electric cars; EV; EVs; charging station; charging stations; charging point; charging points; charging infrastructure; gigafactory; battery plant; battery factory; automotive industry; car industry; carmakers; automakers
Steel & metals (15)	steel; steelmaking; steelworks; blast furnace; blast furnaces; electric arc furnace; direct reduction; iron ore; pig iron; coking coal; aluminium; aluminum; steelmaker; steelmakers; metals industry
Chemicals (18)	chemical plant; chemical plants; petrochemical; petrochemicals; steam cracker; specialty chemicals; speciality chemicals; basic chemicals; fertilizer; fertiliser; fertilizers; fertilisers; industrial gases; chemical recycling; chemical industry; chemicals industry; chemicals sector; chemical makers
Semi-conductors (17)	semiconductor; semiconductors; microchip; microchips; chip plant; chip factory; chip fab; wafer; wafers; microelectronics; integrated circuit; integrated circuits; advanced packaging; lithography; chipmaker; chipmakers; chip industry
Hydrogen (11)	hydrogen; electrolyser; electrolysers; electrolyzer; electrolyzers; electrolysis; fuel cell; fuel cells; green ammonia; hydrogen industry; hydrogen economy
Critical minerals (15)	critical minerals; critical raw materials; rare earth; rare earths; lithium; cobalt; graphite; gallium; germanium; neodymium; permanent magnet; permanent magnets; mining industry; mining sector; mining companies

Notes: Each term is searched as an exact phrase, and the terms of one country–sector pair are joined by OR. The number of terms is given in parentheses. Queries are restricted to the country's outlet list (Table A.2). The United Kingdom, the United States and South Korea use identical queries.

Table A.4. Definitions

Term	Definition
Article	A single press item from one of the listed outlets in the LexisNexis archive; the raw input.
Industrial policy	Government policies that "explicitly target transforming the structure of economic activity in pursuit of a public goal" (adapted from Juhász, Lane and Rodrik, 2023); the definition given to the screening model.

Retained article	An article whose main subject is a specific, current government action aimed at the sector, taken by the tracked country, in an agreement it is party to, or, for EU members, by the EU.
Measure	One concrete government action: one instrument for one group of beneficiaries, as the government decides it. The unit of analysis.
Package	A plan or act that bundles several instruments. Its measures are counted; the package itself is not.
Initiative	A named programme or act that groups measures, such as an EU regulation or a national strategy. A measure is linked to one only when the press names it as part of it; only EU initiatives span countries.
Measure record	A measure as extracted from one article group, before the merge within country and sector.
Canonical measure	The single record that results from merging all reports of the same measure within a country and sector; all counts refer to canonical measures.
Article group	Articles whose short descriptions are similar (cosine similarity of at least 0.85 to the group centroid); measures are first compared and merged within these groups.
Evidence quote	The passage of the article from which a measure was extracted, in the original language.
Core press	Outlets in tiers T1–T3: national general press, national business press and newswires.
Curated set	Measures with at least one core-press article dated on or after 1 January 2026; the sample for cross-country comparison.
Full dataset	All measures from all outlets, including the 2024–25 measures of the ten pairs read from 2024.
Sector	One of six sectors assigned by fixed boundary rules (Section A.5): EVs and automotive, steel and metals, chemicals, semiconductors, hydrogen, critical minerals.
Policy goal	The public objective a measure pursues: green transformation, industrial competitiveness, resilience and security, defence and space, employment and labour, regional development, digital transformation, physical infrastructure, trade agreements, other.
Status	The stage last reported: in discussion (no formal commitment); announced (committed, not yet implemented); active (funds flowing, construction under way or rules in force); cancelled; expired; not stated.
Level of government	The body that decided or announced the measure: EU, national, regional and local, or not stated.
Value	Public money behind a measure as reported in the press (grants, tax reliefs, public loans, guarantees, equity), subject to the exclusion rules in Section A.5.
First reported	The date of a measure's earliest article in the data; a press observation, not the date of decision or enactment.