TECHNICAL WHITE PAPER

The Optional Work Index and the National AI Compute Index

Methodology, scoring functions, and data provenance

optionalwork.com | Engine version 3.1.0 (two-sided geometric mean)

Abstract

This paper specifies two measurement systems and the functions that compute them. The **Optional Work Index** (OWI) scores, from 0 to 100, the degree to which paid work has become economically optional in a country, defined as the conjunction of two conditions: that machines can perform the work (the **supply** side, carrying AI capability, robotics, and labor-market shift) and that people can afford to stop (the **demand** side, carrying economic abundance and wealth distribution). The supply side alone, machines doing the work while people are pushed out, is displacement, not optionality; the index therefore combines the two sides by a geometric mean, which collapses toward zero whenever either side is weak. The **National AI Compute Index** scores effective national AI-processing capacity as raw hardware conditioned by a six-axis enablement factor. A third measure derives per-country workforce displacement from the OWI's own supply-side capability scores applied to occupational structure. Every figure is computed from verifiable public series by fixed, deterministic formulas; missing inputs are omitted, never imputed; the only element that is not bit-reproducible is a bounded news signal. The sections below give the exact scoring functions, thresholds, and sources.

1. Problem statement and construction

The conventional automation metric counts task exposure: the fraction of an occupation a machine could perform. Exposure is necessary but insufficient for the question this index addresses, which is not whether machines can displace labor but whether displacement has become **optional** rather than forced. Optionality requires two independent facts to hold at once. The first is technological, that the productive system no longer depends on a given worker's labor. The second is distributive, that the system transfers enough output to that worker for withdrawal to be survivable. A measure of either fact alone is unfalsifiable against the other: high capability with no distribution is displacement with negative welfare consequences, and broad distribution with weak capability is not optionality either.

The OWI encodes this conjunction structurally. Five sub-indices are partitioned into a supply side and a demand side; each side reduces to a 0 to 100 score; and the headline is the geometric mean of the two side scores, so neither side can compensate for the other. The advancement of AI and robotics is measured **only** on the supply side. The same supply-side capability scores drive the workforce-displacement measure. The compute index is computed independently and is not wired into the OWI, because national capacity and global optionality are distinct quantities on distinct horizons. The remainder of the paper is the formal specification: notation and the sub-index scoring pipeline (Section 2), the supply and demand scoring functions with thresholds and sources (Sections 2.3 to 2.4), side aggregation and the geometric headline with its score

semantics (Sections 2.5 to 2.7), the compute index (Section 3), workforce displacement (Section 4), the cross-index relationship (Section 5), and provenance and reproducibility invariants (Section 6).

2. The Optional Work Index

2.1 Notation

Let the sub-index set be `K = { ai_capability, robotics, labor_displacement, economic_abundance, wealth_distribution }`. Each `k ∈ K` is assigned to a side `side(k) ∈ { supply, demand }` and a within-side weight `w_k` with the side's weights summing to 1. The horizon constants are `PREDICTION_START = 2024-01-01`, `END_10YR = 2034-01-01`, `END_20YR = 2044-01-01`.

Side (0.50 each)	Sub-index (k)	Within-side weight w_k
supply	`ai_capability`	0.40
(can machines do it)	`robotics`	0.31
	`labor_displacement`	0.29
demand	`economic_abundance`	0.60
(can people afford to stop)	`wealth_distribution`	0.40

Within-side weights are the legacy composite weights (`ai 0.28, robo 0.22, labor 0.20, abundance 0.18, wealth 0.12`) renormalized inside each side.

2.2 Sub-index scoring pipeline

Every sub-index score combines a live-data term with three bounded adjustments and is clamped to the unit scale interval:

$$s_k = \mathrm{clip}_{[0,\,100]}\left(L_k + M_k + N_k + d_k\right) \tag{1}$$

Here `L_k` is the live-data term: the arithmetic mean of the present component points for k. Each component `j` maps its public series to points `p_j` on a fixed piecewise scale, given per component in 2.3 and 2.4. Most components are scored 0 to 100, so their mean lies on the 0 to 100 scale:

$$L_k = \frac{1}{|k|}\sum_{j \in k} p_j \tag{2}$$

Averaging over present components keeps the sub-index on the component scale and handles partial data directly: a missing component drops out of the average rather than being scored zero. A sub-index reaches 100 only when every present component reaches 100; the proxy-fed components and the basic-services pillar carry lower maxima and are flagged in 2.3 and 2.4. `M_k` is the milestone bonus from confirmed, structured events, log-scaled and capped:

$$M_k = \min\left(8,\ 2.8\log_2\left(n_k + 1\right)\right) \tag{3}$$

The term `N_k = clamp_[-5,+5]` is the news signal: a fixed prompt scores an hourly snapshot of current headlines for relevance to k and detects milestone confirmations, using model claude-sonnet-4-6; it never emits a score directly and its contribution is bounded to five points. `d_k ∈ [-0.8,+0.8]` is a deterministic daily micro-signal seeded from the date and the key. The scoring formula is deterministic for a fixed input set: the same series values, confirmed-milestone set, and headline snapshot reproduce the same score, and a milestone counts only when it matches a confirmed entry in the fixed milestone list after URL and title deduplication. The news term alone is not bit-reproducible across cold runs, because the headline snapshot changes hour to hour; the model call is pinned to temperature 0 to remove model-side variance, and the term stays within its five-point bound. If k has no live components and no confirmed milestones, `s_k = 0`.

2.3 Supply-side scoring functions

The supply score reads whether the economy still requires human labor. Each component below maps a public series to points by a fixed piecewise rule; the sub-index score `L_k` is the mean of the present components (Section 2.2).

Sub-index	**Component (source)**	**Points rule**
ai_capability	ai_frontier (global frontier benchmark, 0..100)	pass-through: p = clamp(value); identical for every country (capability ceiling)
	ai_adoption (FRED productivity YoY ρ, or internet-user % u)	ρ: ≥30→100; ≥20→88; ≥12→75; ≥8→55+; ≥5→38+; ≥3→22+; ≥2→12+; ≥1→5+; else ~5ρ. u-proxy: 95→26, 90→20, 75→14, 50→8, 25→4 (proxy ceiling 26)
robotics	robot_density (IFR per 10k d, or mfg VA % GDP)	d: ≥2500→100; ≥1500→85; ≥1000→70; ≥600→50+; ≥300→32+; ≥150→18+; ≥50→8+. VA-proxy: 30→24, 20→16, 10→8 (proxy ceiling 24)
	automation_gap g = output growth − employment growth (FRED IPMAN vs MANEMP)	g: ≥35→100; ≥22→85; ≥15→70; ≥10→50+; ≥6→28+; ≥3→15+; ≥1→8+; ≥0→5+
labor_displacement	unemployment r (BLS / WB-ILO)	r: ≥40→100; ≥25→70; ≥15→40; ≥10→22; ≥8→15; ≥6→8; ≥5→5; ≥4→2; else 0
	productivity YoY ρ (FRED / WB)	ρ: ≥25→100; ≥15→70; ≥8→38; ≥5→25; ≥3→16; ≥2→10; ≥1→5; else 4ρ
	job_openings (FRED JTSJOL) or emp/pop slack	openings o = % below 12,000K peak: o≤40→round(0.4o); o>40→min(100, 16+1.4(o−40)). emp/pop e: ≤15→100, ≤25→70, ≤35→45, ≤40→20, ≤50→12, ≤55→6, ≤60→3; else 0

The supply side rises with frontier capability, robot deployment, the divergence of output from employment (machines producing more with fewer workers), and labor slack consistent with displacement. Each component is scored on its own piecewise scale and the sub-index is the simple mean of its present components. Most components reach 100 only at conditions far beyond any country today (sustained extraordinary productivity, mass robot density, displacement-scale unemployment or slack). Two proxies used only

where the primary series is absent carry lower maxima, the internet-user adoption proxy at 26 and the manufacturing value-added robotics proxy at 24, and are flagged as weak.

2.4 Demand-side scoring functions

The demand score reads whether people can afford to stop. Economic abundance is a five-pillar reading of the cost of necessities (higher score means necessities are cheaper or more accessible, so work is less required); wealth distribution reads the income floor and how broadly gains are shared. As above, `L_k` is the mean of the present components (Section 2.2).

Sub-index	**Component (source)**	**Points rule**
economic_abundance	healthcare: out-of-pocket % of health spend (WB SH.XPD.OOPC), lower better	≤0.5→100; ≤2→88; ≤4→72; ≤7→55; ≤10→40; ≤20→28; ≤35→16; ≤50→8; ≤65→3; else 0
	education: public spend % GDP (WB SE.XPD.TOTL), higher better	≥14→100; ≥12→80; ≥10→62; ≥8→45; ≥7→30; ≥5→20; ≥4→12; ≥3→7; ≥2→3; else round(s)
	food: real per-capita consumption growth (WB proxy)	≥10→100; ≥8→70; ≥6→45; ≥5→24; ≥3→16; ≥1→8; ≥0→4; else 0
	basic services: 0.30 electricity + 0.25 water + 0.25 sanitation availability (WB) + 0.20 broadband affordability (ITU)	weighted composite c over present parts; broadband basket as % GNI: ≤2→100, ≤5→75, ≤10→50, ≤20→25, else 10. p = round(c · 22/100), pillar maximum 22
	housing: GDP-per-capita anchor (weakest pillar)	≥160k→100; ≥120k→70; ≥80k→40; ≥60k→14; ≥40k→10; ≥20k→6; ≥8k→3; else 1
wealth_distribution	income_floor: social-protection coverage % (WB / ILO), higher better	≥99.9→100; ≥99→85; ≥95→65; ≥90→44; ≥70→32; ≥50→20; ≥30→10; ≥15→5; else round(0.2c)
	inequality: Gini (WB SI.POV.GINI), lower better	≤15→100; ≤18→78; ≤22→55; ≤25→30; ≤30→22; ≤35→14; ≤45→7; ≤55→3; else 0
	labor_income_share: labor share of GDP, or US real-wage YoY (ILO / FRED)	share: ≥70→100; ≥67→80; ≥63→55; ≥60→26; ≥52→18; ≥45→11; ≥38→6; else round(0.1s). US real-wage YoY: ≥15→100; ≥12→80; ≥10→55; ≥8→26; ≥5→18; ≥3→11; ≥2→7; ≥1→4; ≥0→round(3y)

Economic abundance and wealth distribution are each the simple mean of their present components. Most pillars reach 100 at their strongest tier; basic services is deliberately held to a pillar maximum of 22, and housing and food use anchor proxies, the weakest pillars by data quality and first in line for replacement. Healthcare and education weigh equally with the others in the mean but matter most in interpretation, since medical and tuition burdens most directly force continued wage labor and their data are the strongest. The basic-services pillar combines utility availability (electricity, water, sanitation from the World Bank) with broadband affordability (the ITU fixed-broadband

price basket as a share of GNI per capita), weighted 0.30, 0.25, 0.25 and 0.20 over the parts present, the composite then scaled to the pillar maximum of 22. A component with no data is dropped from the mean, not set to zero.

2.5 Side aggregation and the geometric headline

Each side score is the within-side weighted mean over the sub-indices present (so a missing sub-index drops out of both numerator and denominator):

$$\text{Score}_{\text{side}} = \frac{\sum_{k \in \text{side}} s_k \, w_k}{\sum_{k \in \text{side}} w_k} \tag{4}$$

The headline is the clamped geometric mean of the two side scores:

$$\text{OWI} = \min\left(100,\ \sqrt{\text{Score}_{\text{supply}} \cdot \text{Score}_{\text{demand}}}\right) \tag{5}$$

The geometric mean is chosen because optionality is a conjunction. It has three properties an arithmetic mean lacks: it never exceeds the arithmetic mean, so it cannot be more optimistic; it lies between the weaker and the stronger side, so the weaker side anchors the headline; and it equals zero whenever either side is zero. Concretely, `Score_supply = 60, Score_demand = 10` gives `√600 ≈ 24.5`, against an arithmetic 35; raising demand to 60 gives `√3600 = 60`, where the two means coincide. In the current period the demand side sits low almost everywhere, so an arithmetic mean would let supply-side capability inflate the index and report displacement conditions as optionality. The geometric mean prevents this by construction.

Stated formally, for side scores `S, D ∈ [0,100]`:

$$\text{(P1)} \quad \min(S, D) \ \le\ \sqrt{S \cdot D}\ \le\ \max(S, D)$$

$$\text{(P2)} \quad \sqrt{S \cdot D}\ \le\ \frac{S + D}{2}, \qquad \text{equality iff } S = D$$

$$\text{(P3)} \quad \sqrt{S \cdot D} = 0\ \Leftrightarrow\ S = 0 \text{ or } D = 0$$

(P1) and (P3) follow directly from the definition. (P2) is the arithmetic-geometric mean inequality: since `(√S − √D)² ≥ 0` expands to `S + D ≥ 2√(S·D)`. (P2) guarantees the headline is never more permissive than the arithmetic average, and the equality condition shows the two coincide only when the sides are equal; for every unequal pair the geometric mean is strictly lower, by an amount that grows with the divergence between the sides.

2.6 Momentum, trajectory, and time progress (not in the headline)

Momentum is computed but deliberately excluded from the headline. It is a trailing-confirmation factor

$$\mu = 1 + \min(0.15,\ 0.025\, c) \tag{6}$$

and is published only as a separate trajectory readout. Time progress is the elapsed fraction of each horizon, `P10 = 100 · elapsed / (END_10YR − START)` and similarly `P20` to 2044; trajectory status compares the ratio `OWI / P10` against fixed bands. None of these feed `OWI`; the headline is the geometric mean and nothing else.

ratio = OWI / P10	**Status**	**delta**
≥ 1.25	Accelerating	+
≥ 1.08	Ahead of schedule	+
≥ 0.88	On track	~
≥ 0.72	Lagging	-
< 0.72	Behind schedule	-

The daily micro-signal `d_k` is itself deterministic: it is seeded from the date and the sub-index key (a sum of three uniform draws divided by 300, scaled by 0.8), so it is identical on re-run for a given day and contributes at most 0.8 points.

2.7 Score semantics and the ceiling

Because `OWI = √(Score_supply · Score_demand)`, the value 100 is attainable only when both side scores equal 100, which in turn requires every present sub-index to reach its own maximum, and the component thresholds place those maxima at conditions far outside historical norms (sustained extraordinary productivity growth, mass robot deployment, near-universal social-protection coverage with low inequality). A score of 100 is therefore an arithmetic ceiling, not a forecast and not a claim that work is optional. Intermediate values are monotone in both sides: the index can rise only as capability and affordability rise together.

2.8 Algorithm and a worked example

The complete per-country computation is the following deterministic procedure:

```
for k in K:
    L = sum(points_present) / count(points_present)   # arithmetic mean of present
component points
    M = min( 8, 2.8 * log2(n_k + 1) )    # confirmed-milestone bonus
    N = clamp( news_k, -5, +5 )          # bounded headline signal
    d = daily(date, k)                   # deterministic, |d| <= 0.8
    s[k] = clamp( L + M + N + d, 0, 100 )
for side in {supply, demand}:
    Score[side] = sum( s[k]*w[k] ) / sum( w[k] )   # present k in side
OWI = min( 100, sqrt( Score[supply] * Score[demand] ) )
```

A worked instance with illustrative inputs (not a real country's published figures). For `ai_capability`, suppose the frontier benchmark scores 62 and the adoption term scores 20, both on the 0 to 100 component scale. Then `L = (62 + 20)/2 = 41.0`. With `n = 4` confirmed milestones, `M = min(8, 2.8·log2 5) = 6.50`; news `N = +2.0`; daily `d = +0.3`. So `s[ai_capability] = clip(41.0 + 6.50 + 2.0 + 0.3) = 49.80`.

Taking `s[robotics] = 21.0` and `s[labor_displacement] = 13.5`, the supply side is

$$\text{Score}_{\text{supply}} = 49.80 \cdot 0.40 + 21.0 \cdot 0.31 + 13.5 \cdot 0.29 = 30.35$$

Taking `s[economic_abundance] = 17.0` and `s[wealth_distribution] = 26.0`, the demand side is `17.0·0.60 + 26.0·0.40 = 20.60`, and the headline is

$$\mathrm{OWI} = \sqrt{30.35 \cdot 20.60} = \sqrt{625.2} \approx 25.00$$

The arithmetic mean of the two sides would be 25.48; the geometric mean is lower, and the gap widens as the sides diverge. Had demand been 8.0 instead of 20.60, the headline would be `√(30.35·8) ≈ 15.58` against an arithmetic 19.18, so the stronger supply side cannot offset the weaker demand side.

2.9 Coverage and partial data

Because a missing sub-index is dropped from its side's weighted mean rather than scored zero, a side can be computed from fewer than its full complement of sub-indices. The engine therefore reports coverage alongside the score: for each side it returns the count of sub-indices that had real data against the side's total (`supply` has three, `demand` has two). A side score resting on partial coverage is still a well-defined weighted mean over the present sub-indices, but the coverage counts make the difference visible, so a high score backed by one present sub-index is not silently treated as equivalent to one backed by all of them. This is the no-imputation rule applied at the level of the side rather than the component.

2.8 Threshold Selection

The scoring tables in 2.3 and 2.4 fix where each component earns its points. Those breakpoints are not arbitrary, but neither are they all of one kind. They come from four distinct sources, and the engine is explicit about which applies where, because a threshold defended as an observed fact carries different weight from one chosen by judgment.

Observed distribution and historical extreme. Most tiers are placed against the range of values countries actually exhibit, with the middle tiers sitting where real national data concentrates so that the score discriminates across the realistic range rather than saturating. The top tiers added above today's best, productivity year-on-year at or above 25 percent, robot density at or above 2,500 per ten thousand, out-of-pocket health spending at or below 0.5 percent, a Gini at or below 15, are deliberate extrapolations beyond any country today. They exist so the scale can reach 100 in principle; none is a prediction that any country will arrive there.

Policy benchmark. A few thresholds are anchored to recognised policy references. Public education spending uses the 4 to 6 percent of GDP range familiar from international education targets as its discriminating band; the 14 percent tier that scores 100 is a theoretical ceiling far above any sustained national budget, set precisely so that no country reaches 100 on this pillar from current data. Social-protection coverage is scored against the universal-coverage goal, so that near-universal coverage approaches 100.

Normative judgment. Some choices are openly normative: the milestone bonus is capped at 8 and the news signal is bounded to plus or minus 5 so that neither can overwhelm the measured series; the basic-services pillar is held to a maximum of 22; and the components inside each sub-index are combined by a simple mean rather than by fixed unequal weights. These are design decisions, and the sensitivity analysis in 2.10 quantifies how much they matter.

Three specific choices the thresholds raise are worth answering directly. **Why does 14 percent education spending score 100?** Because no country sustains 14 percent of GDP on public education; the tier is a ceiling, not a target, chosen so the pillar reaches 100 only in the limit while the realistic 3 to 8 percent range stays in the discriminating middle of the scale. **Why is basic services capped at 22?** In developed economies utilities are near-universal, so at the top the pillar carries little discriminating signal; holding it to 22 prevents a near-saturated pillar from dominating the abundance mean, and the cap is flagged as interim, to be revisited once the broadband-affordability (ITU) component is populated across countries. **Why average the abundance pillars equally?** Because there is no defensible empirical basis for a fixed unequal weighting across healthcare, education, food, basic services, and housing affordability. Equal weight is the neutral prior. The engine deliberately does not encode the intuition that healthcare and education matter most as a numeric weight, since that would smuggle a normative ranking into the score; instead, relative importance is expressed through tier design, how quickly each pillar rises with improving conditions, which is visible and adjustable, rather than through hidden weights.

All thresholds are fixed and versioned. A score computed under one set of tiers is stamped with the model version, so changing a threshold invalidates prior results rather than silently re-scaling them.

2.9 A worked example on real data

The illustrative instance in 2.7 used invented inputs. The table below instead shows seven real countries computed by the engine from current public series, chosen to span the range: the United States, a capability leader with comparatively higher demand-side affordability; China, a manufacturing and robotics leader; Ghana and South Africa, two African economies at different income levels; the European Union aggregate; Brazil, a middle-income economy; and Haiti, a low-income economy weak on both sides. Supply is the weighted mean of the three supply sub-indices, demand the weighted mean of the two demand sub-indices, and OWI their geometric mean. Sub-index scores are shown in parentheses.

Country	**Supply (ai / robotics / labor)**	**Demand (abundance / wealth)**	**OWI**
USA	29.45 (38.8 / 28.7 / 17.4)	24.10 (25.9 / 21.4)	26.64
China	38.14 (38.8 / 64.2 / 9.4)	19.03 (20.9 / 16.2)	26.94
Ghana	19.66 (32.8 / 9.2 / 12.7)	13.73 (15.4 / 11.2)	16.43
South Africa	29.27 (35.8 / 9.2 / 41.7)	20.33 (27.4 / 9.7)	24.39
EU	18.32 (29.3 / 8.8 / 13.3)	15.37 (20.4 / 7.8)	16.78
Brazil	21.19 (35.8 / 9.2 / 13.9)	17.98 (22.2 / 11.7)	19.52
Haiti	21.97 (30.8 / 10.2 / 22.4)	10.48 (9.7 / 11.7)	15.17

The comparison makes the geometric mean concrete. China's supply side (38.14) is well above the United States' (29.45), driven by far higher robot density, yet the two headline scores are almost identical (26.94 against 26.64), because China's demand side is lower (19.03 against 24.10). An arithmetic mean would have ranked China materially higher and reported its stronger automation capability as greater optionality; the geometric mean holds the two countries together, because optionality requires both that machines can do the work and that people can afford to stop, and China leads on only the first. Ghana sits low on both sides (16.43). South Africa scores higher (24.39), but largely because its very high real unemployment, around a third of the labor force, lifts the labor-displacement proxy on the supply side; this is a reminder that unemployment is an imperfect stand-in for automation-driven displacement, and the headline should be read with that in mind. The European Union aggregate sits at 16.78 on a weaker measured demand side, Brazil falls in the middle (19.52), and Haiti is low on both sides, so neither the arithmetic nor the geometric mean rescues it (15.17). For all three, the basic-services pillar has no data in the current snapshot and is dropped from the abundance mean rather than imputed, so abundance here is the mean of the four pillars present.

2.10 Sensitivity analysis

A score that moved sharply under small changes to its weights or that depended on its weakest proxies would be hard to trust. Two perturbations test this on the same three countries. The first varies every within-side weight by plus or minus 10 percent,

renormalizing inside each side, and reports the resulting range of the headline. The second drops the weakest proxy inputs, the food and housing affordability pillars on the demand side and the adoption proxy on the supply side, and recomputes from the remaining components.

Country	**Baseline OWI**	**Weights moved +/- 10%**	**Weak proxies excluded**
USA	26.64	26.46 to 26.80	30.79
China	26.94	26.64 to 27.24	28.00
Ghana	16.43	16.20 to 16.64	20.07
South Africa	24.39	24.12 to 24.66	33.10
EU	16.78	16.57 to 16.98	23.05
Brazil	19.52	19.24 to 19.78	23.45
Haiti	15.17	15.04 to 15.30	17.98

Under weight perturbation the headline is highly stable: the widest band is China's, and even there the score moves by less than 0.6 points, because the geometric mean of two sides is far less sensitive to within-side reweighting than the side scores themselves. The index is therefore not an artefact of the particular weights chosen. Proxy exclusion tells a more pointed story. Dropping the weakest proxies, the food and housing affordability pillars on the demand side and the adoption proxy on the supply side, raises the headline rather than lowering it, by under 0.2 points for the United States but by roughly 1 to 9 points elsewhere (China to 28.00, Ghana to 20.07, South Africa to 33.10, the EU to 23.05, Brazil to 23.45, Haiti to 17.98). The direction is informative: for most countries these weak pillars currently score low and drag the measured score down, so the present index is conservative where proxy data is poor, and better feeds would tend to raise scores rather than lower them. Either way the dependence is real and concentrated in data-poor settings, which is exactly why those pillars are flagged in 2.4 and are first in line for replacement. The sensitivity analysis is reported so this dependence is visible rather than hidden.

3. The National AI Compute Index

The compute index scores effective national AI-processing capacity: raw hardware conditioned by whether it is usable. It is standalone and does not enter the OWI. The present version is explicitly a tiered proxy model: hardware and several enablement axes are scored as curated 0 to 3 tiers rather than measured directly, standing in for concrete quantities, installed accelerator counts and H100-equivalents, datacenter power capacity, in-country cloud-region availability, export-control exposure, and the split between training and inference capacity, that later versions are intended to measure.

Limitation. This is the least authoritative of the three measures, and deliberately flagged as such. The hardware tiers and the control, talent, and software enablement axes are curated 0 to 3 proxy values, assigned by judgment against public information, not measured series. A reader can reasonably contest a particular country’s tier, and two careful analysts could assign different tiers from the same evidence. Until the curation rules are published in full or, better, the tiers are replaced by measured quantities (installed accelerator counts and H100-equivalents, datacenter power capacity, in-

country cloud regions, and the training-to-inference split), the compute scores should be read as indicative rankings rather than precise measurements, and they are kept out of the OWI for exactly this reason.

3.1 Raw, enablement, effective

$$\text{Raw} = 100\left(0.70 \cdot \frac{\min(\text{accel}, 3)}{3} + 0.30 \cdot \frac{\min(\text{dc}, 3)}{3}\right) \quad (7)$$

where `accel` and `dc` are accelerator/cluster and data-center capacity tiers (0..3). Enablement is the geometric mean over the six axes that have a value, expressed as a factor in [0,1]:

$$\text{Enablement} = \sqrt[n]{\prod_{i=1}^{n} \frac{a_i}{100}} \quad (8)$$

$$\text{Effective} = \text{Raw} \cdot \text{Enablement} \quad (9)$$

Effective is the published headline. Idle or unreachable hardware scores low because high `Raw` times a low enablement factor yields a low `Effective`. A verdict band is attached: `≥60 Strong, ≥35 Moderate, ≥15 Limited, else Minimal`.

3.2 Enablement axes

Axis a_i (0..100)	Definition
accessible	100 · (0.70 · min(regions,8)/8 + 0.30 · prog). regions = hyperscaler cloud regions in-country, prog = national public-compute program {0,1}. Not floored: no in-country access scores 0.
affordable	100 · ($\log_{10}$ g − $\log_{10}$ 1500) / ($\log_{10}$ 75000 − $\log_{10}$ 1500), g = GNI per capita at PPP; null if g ≤ 0.
reliable	E · κ(supply), E = electricity-access %, κ = {1:1.0, 2:0.7, 3:0.3} continuity factor for advanced-chip supply / export-control exposure.
control / talent / software	tier floor φ(tier), φ = {0:20, 1:50, 2:75, 3:100}. These three are coarse curated tiers, the softest inputs, designed to be replaced by published data.

The geometric mean across axes is deliberate: a single collapsed axis (no access, no power) should pull effective capacity toward zero, which an average would not do. The informative quantity is the gap between `Raw` and `Effective`: hardware that exists but cannot be reliably used registers as a large gap. As an arithmetic illustration, `Raw = 100` with enablement 0.7 yields `Effective = 70`; the same hardware at enablement 0.4 yields 40.

3.3 Provenance principle: capacity is a policy choice

National AI compute is decoupled from GDP: a lower-income country with an active sovereign program can hold real usable capacity while a richer country without one may not. The index therefore rejects development-tier guesses for uncurated countries, since such guesses encode the false assumption that capacity tracks income, and instead uses real per-country inputs that authoritative national sources can supersede.

3.4 Worked example

Illustrative inputs. Accelerator tier `accel = 3`, data-center tier `dc = 2`: `Raw = 100·(0.70·1.0 + 0.30·0.667) = 90.0`. Axes: `accessible` with 4 regions and a national program is `100·(0.70·0.5 + 0.30·1) = 65.0`; `affordable` at `g = 18000` is `100·(log₁10 18000 − log₁10 1500)/(log₁10 75000 − log₁10 1500) ≈ 63.5`; `reliable` at 92% electricity and supply tier 2 is `92·0.7 = 64.4`; `control = 75` (tier 2), `talent = 50`, `software = 50` (tier 1).

The enablement factor is the geometric mean of the six axes as fractions:

$$\text{Enablement} = (0.650 \cdot 0.635 \cdot 0.644 \cdot 0.750 \cdot 0.500 \cdot 0.500)^{1/6} \approx 0.607$$

so `Effective = Raw · Enablement = 90.0 · 0.607 ≈ 54.6`, verdict Moderate. The raw-to-effective gap of roughly 35 points is the signal: the hardware exists, but talent, software, and supply continuity hold usable capacity well below the hardware ceiling. If a required axis (for example affordability when `g ≤ 0`) has no value, it is omitted from the geometric mean rather than set to zero.

Effective	Verdict	Reading
≥ 60	Strong	usable capacity at scale
35 to 60	Moderate	real but constrained capacity
15 to 35	Limited	thin or heavily conditioned
< 15	Minimal	little usable national capacity

4. Workforce Displacement

Displacement is derived from the OWI's own supply-side capability scores, not a separate model. A susceptibility ceiling `σ(g, channel)` gives, for each ISCO-08 major group g and channel ∈ {ai, robotics}, the share of the group's work that is structurally the kind a machine could ever do. Ceilings are slow-moving properties of the occupation:

Group	Label	σ ai	σ robotics	Driver
1	Managers	0.40	0.10	
2	Professionals	0.55	0.10	cognitive
3	Technicians	0.50	0.35	
4	Clerical support	0.80	0.30	highest AI
5	Service and sales	0.40	0.45	
6	Skilled agricultural	0.15	0.65	physical
7	Craft and trades	0.20	0.70	physical
8	Plant / machine operators	0.20	0.85	highest robotics
9	Elementary occupations	0.15	0.75	physical
0	Armed forces	0.20	0.30	

How much of that susceptible work is replaceable now is set by the OWI capability score for the channel:

$$\text{replaceable}(g, c) = \sigma(g, c) \cdot \frac{\text{cap}_c}{100} \tag{10}$$

$$\text{combined}(g) = 1 - (1 - r_{\text{ai}}(g))(1 - r_{\text{rob}}(g)) \tag{11}$$

where `cap_ai` and `cap_robotics` are the OWI's `ai_capability` and `robotics` sub-index scores. The national share for a channel is the employment-weighted sum over groups, with the occupation mix taken from ILOSTAT (ISCO-08), available by sex and age segment:

$$D_c = \sum_g \text{share}(g) \cdot \text{replaceable}(g, c) \tag{12}$$

Because the multiplier is the OWI capability score and the occupation mix moves slowly, the displacement figure tracks the same supply-side signals that drive the index. It estimates **exposure** (an upper bound on what current capability could replace), not realised displacement, whose timing depends on adoption, cost, and regulation that the measure does not model. A segment with no occupational data is omitted, never synthesised.

Worked instance. With OWI capability scores `cap_ai = 52` and `cap_robotics = 21`, clerical support (group 4, `σ ai = 0.80, σ robotics = 0.30`) has `replaceable(4, ai) = 0.80·0.52 = 0.416` and `replaceable(4, robotics) = 0.30·0.21 = 0.063`, so `combined(4) = 1 − (1−0.416)(1−0.063) = 0.453`: about 45% of clerical work is replaceable now at this capability level. If clerical is 12% of national employment, its contribution to the combined national share is `0.12 · 0.453 = 0.054`, and the national figure is the sum of such contributions across all groups.

The same computation runs over employment segments. Where ILOSTAT provides the occupation mix by sex and by age, the engine produces segmented displacement shares as well as the national figure, applying the identical σ ceilings and the identical OWI capability multipliers to each segment's mix. Differences between segments therefore come only from differences in occupational composition, not from any change in the model, which keeps the segmented figures comparable to one another and to the national total.

5. Relationship between the indices

Compute is the upstream input to the OWI's supply side: the capability the OWI reads as advancing must be trained and served on the hardware the compute index measures. The two are nonetheless kept arithmetically separate. Global capability is increasingly reachable through cloud access and open models regardless of a country's own compute score, so low national compute does not mean a country is untouched by the optional-work transition, only that it has less control over the capability shaping it. Coupling the indices would conflate the global condition (optionality) with the national endowment (effective compute), two quantities on different horizons. They are reported together and computed apart.

6. Data provenance and reproducibility

Three invariants hold across both indices. **No fabrication:** every component is sourced from a named public series (FRED, BLS, World Bank indicator codes, IFR, ILOSTAT), and a component with no data is dropped from the component mean rather than imputed, so a score reflects only data that exists. **Determinism:** structure, weights, thresholds, and formulas are fixed, so identical inputs reproduce an identical score; the language-model layer is bounded to ±5 points and confined to signal extraction from public text. **Supersession:** curated inputs can be replaced by authoritative national figures, and a result computed under a superseded model version is invalidated by a schema stamp rather than served stale. These are properties of the method and the data, independent of any particular deployment.

7. Limitations and conclusion

The limitations are specific and named in the engine. Several inputs are proxies whose signal is indirect: AI adoption outside the United States is approximated by internet-user share; housing affordability by GDP per capita; food affordability by per-capita consumption growth; and labor displacement partly by unemployment, which also moves with the business cycle, demographics, and external shocks, so the labor sub-index combines it with productivity and demand slack rather than reading it alone. On the compute index the sovereign-control, talent, and software axes are coarse tiers awaiting published data. The OWI is anchored to one named prediction and the horizons 2034 and 2044 from a 2024 start; a different anchor reframes the presentation without changing the data. The ceiling at 100 remains an arithmetic bound, reached only under conditions far outside historical norms, and the displacement measure is exposure, not realised displacement.

The construction yields a falsifiable and reproducible reading of a question usually addressed only by assertion. Supply and demand are scored separately and combined by a geometric mean that no strong side can offset; raw compute is distinguished from usable compute; and displacement is derived from the measured capability scores rather than an independent estimate. Every published figure is computed from public series by a fixed function, is traceable to its source, and is recomputed deterministically as those series are updated.

Appendix A. Consolidated data sources

Series / source	Indicator	Feeds
Global frontier benchmark	Composite capability 0..100	`ai_capability`
FRED OPHNFB	Labor productivity YoY	`ai_capability`, `labor_displacement`
World Bank IT.NET.USER.ZS	Internet users % (adoption proxy)	`ai_capability`
IFR	Robot density per 10k workers	`robotics`
FRED IPMAN, MANEMP	Output vs employment (automation gap)	`robotics`
BLS / World Bank ILO	Unemployment rate	`labor_displacement`

FRED JTSJOL	Job openings	`labor_displacement`
World Bank SH.XPD.OOPC	Out-of-pocket health spend %	`economic_abundance`
World Bank SE.XPD.TOTL	Public education spend % GDP	`economic_abundance`
World Bank	Per-capita consumption growth; GDP per capita	`economic_abundance`
World Bank	Electricity, basic water, basic sanitation access %	`economic_abundance (basic services)`
ITU	Fixed-broadband price basket % GNI per capita	`economic_abundance (basic services)`
World Bank / ILO	Social-protection coverage %	`wealth_distribution`
World Bank SI.POV.GINI	Gini coefficient	`wealth_distribution`
ILO / FRED	Labor income share; real wages	`wealth_distribution`
World Bank	GNI per capita at PPP	`compute: affordable`
World Bank	Electricity access %	`compute: reliable`
ILOSTAT	ISCO-08 employment by group / sex / age	`displacement: occupation mix`

Conventions. World Bank indicators are read at their most recent available value within a two-year window; FRED series enter as year-over-year change except where a level is specified; the global frontier benchmark is identical across countries and represents the capability ceiling, while every other series is per-country. Tier inputs (`accel, dc, control, talent, software, supply` and the compute program flag) are curated 0 to 3 values. Any series may be superseded by an authoritative national figure, and any series with no value is omitted from its component sum rather than imputed, so coverage rather than a placeholder records the absence.

Appendix B. Notation and symbol glossary

Symbol	Meaning
K	The five sub-indices: ai_capability, robotics, labor_displacement (supply); economic_abundance, wealth_distribution (demand).
side(k)	The side a sub-index belongs to, supply or demand.
w_k	Within-side weight of sub-index k; weights within each side sum to 1 (ai 0.28, robotics 0.22, labor 0.20, abundance 0.18, wealth 0.12, renormalized inside each side).
L_k	Sub-index score: the mean of the present components of k, on the 0 to 100 scale; a missing component drops out of the mean.
M_k	Milestone bonus from confirmed structured events, log-scaled and capped at 8 points.
d_k	Daily deterministic micro-signal, seeded from the date and the sub-index key, bounded to 0.8 points.
Score_supply, Score_demand	Side scores: the within-side weighted mean over the sub-indices present on each side.
OWI	The headline: the clamped geometric mean of the two side scores, sqrt(Score_supply * Score_demand).
raw	Compute index raw hardware capacity, from accelerator/cluster and data-center tiers (0..3).
enablement	Compute index enablement factor in [0,1]: the geometric mean over the six enablement axes present.
effective	Compute index effective capacity: effective = raw * enablement.
PREDICTION_START	Horizon origin, 2024-01-01.
END_10YR, END_20YR	Horizon endpoints, 2034-01-01 and 2044-01-01.
P10, P20	Time-progress fractions of each horizon; reported as a trajectory readout only, not in the OWI.