ORC VALUE STANDARD 0.2

The ORC Value Standard: a unit of measure and a business model for subscription pricing

Published by PayOrBuild on 13 November 2026. The method that every label on payorbuild.com applies.

Summary

A subscription's price can be read against one public number: what it would cost to rebuild the service lawfully. This standard turns that number into a unit, gives it an error range, prices professional finesse into it, lets the owner's real code inform it under seal, and issues a Value Score only to plans that stay cheaper than replication three years ahead.

Part	What it is	Why it holds
The unit	P/R: outlay divided by replication price at standard conditions, read today, 36 months ahead and at the floor	One denominator for every pricing model, as the annual percentage rate is for credit
The finesse hurdle	Know-how counted in decision points: derivable, discoverable or hidden	Discovery gets cheaper with AI. Proof of quality takes calendar time
The seal	An open instrument measures the real code. Only bands, ranges and a hash are published	Audits and security certifications already split evidence this way
The error range	The median with the 10th and 90th percentiles of a seeded simulation	Anyone can recompute it, and its hit rate is published
The Value Score	The cautious discount to replication price at month 36, issued at 90% confidence	A seal that survives the pace of AI means something
The business model	Method, instrument and Index are free. Attestation, the mark, the register feed and buyer tools are paid	It sells only what the framework says cannot be replicated

The example in three lines. COTBase Premium costs about a tenth of its replication price today. On public evidence that advantage holds for 27 months at 90% confidence. No Value Score is issued yet: the plan is still cheaper at month 36 in 68% of outcomes, and the test requires 90%.

What would earn the score. A price at or below about $9 a month, or sealed proof of a component whose quality takes time to verify, or a service that a replica must keep paying for.

Status. This is version 0.2, for comment. The Index and the instrument do not exist yet, and the example uses declared ranges.

What version 0.2 adds

Version 0.2 turns the rule of version 0.1 into a measuring standard: one unit, an error range on every figure, a priced hurdle for professional finesse, a way to measure the owner's real code without publishing it, and a score that looks three years ahead.

Element	Version 0.1	Version 0.2
Headline figure	Replication multiple, as a low and a high case	P/R: price divided by replication price, with an 80% range. The multiple is its inverse
Quality	Fidelity levels F1 to F4: scope, service level, currency	Adds quality parity: the replica must perform as well as the original on a declared public measure
Know-how	Named as a limit: size is not value	Priced as the finesse hurdle, sized in decision points
Evidence	Public only, graded A to C	Four assurance levels. The owner may let an open instrument measure the real code under seal
Time	A parity runway	A P/R curve: spot, forward and terminal
Result	Subscribe, or Premium	Adds the Value Score, issued only when the plan stays below replication price three years ahead, at the cautious end of the range
Organisation	A neutral keeper	A business model: who pays whom, for what, and why the number stays honest

Everything else stands. The standard still prices the lawful route only, advises in one direction only, tests three replicators, separates six layers, reports the floor, and hard-codes no cost.

The unit: price-to-replication

The unit is P/R: what a buyer pays, divided by what it would cost to replicate the same thing lawfully. At 1 the two are equal. Below 1, subscribing is cheaper.

$$\mathrm{P/R}(t) \;=\; \frac{S}{\mathrm{ORC}_t(F, Q, k, T)}, \qquad S = k\,p\,A_T$$

ORC is the price to reverse engineer, as version 0.1 defines it, now stated at a fidelity F and a quality Q. S is the outlay of k users at price p over the horizon. P/R is the inverse of the version 0.1 replication multiple.

Why a unit. Markets compare well when they share one denominator. US lenders have had to disclose the annual percentage rate since the Truth in Lending Act of 1968, so that borrowers can compare credit terms. EU shops must show a unit price beside the selling price to make prices comparable. Investors read a share price against earnings, today and forward. P/R reads a subscription against its replication price in the same way.

Three readings. One formula is read at three dates.

Reading	Replication is priced at	Question it answers
Spot P/R	Today's Index	Is subscribing cheaper than building now?
Forward P/R	The Index trend, 36 months ahead	Will it still be cheaper after three more years of AI progress?
Terminal P/R	The floor: tokens and oversight free	Is the price protected by something AI does not make cheap?

Standard conditions. Two P/R figures can be compared only if they were measured the same way.

Condition	Standard value
Buyer	The plan's standard profile: one user and the capabilities that profile uses. Per-seat plans are also quoted at 10 and 100 seats
Fidelity and quality	F1 core at quality parity for one user. F4 living service for groups and entrants
Horizon	36 months, no discounting
Price	The plan's lowest list price for a commitment of 12 months or less
Wage	The median wage of the skill the replica needs, from public statistics
Index	The current epoch. Forward readings use its published trend
Statistic	The median, with the 10th and 90th percentiles of a seeded simulation

Every pricing model converts. Only the outlay changes.

Pricing model	Outlay over the horizon	Note
Flat subscription	Price times months	
Per seat	Price times seats times months	A replica's cost is mostly fixed, so P/R rises with seats. The label quotes the seat count at parity
Usage-based	Rate times the standard profile's usage	The replica's own running cost per unit goes into ORC
Tiers and add-ons	The price step between tiers	Set against the ORC of what the step adds
Freemium	Zero for the free tier	The paid tier is assessed as a step
One-off licence	Price plus maintenance	
Commission or transaction fee	Fee times the standard profile's volume	
Bundle	Each plan and each step separately	

Pricing model	Outlay over the horizon	Note
Ad-funded	None	Out of scope: the buyer pays in attention

A plan at P/R 0.10 costs a tenth of its replication price. A plan at 1.3 carries a 30% premium over replication, and the label says what the premium buys.

The finesse hurdle

Professional finesse is what a working copy lacks: the gap between doing the same things and doing them as well. Version 0.2 prices that gap.

What economics says. The gap is what has always protected creators. Teece showed that who profits from an innovation depends on how tacit its knowledge is and on who owns the complementary assets. Von Hippel called information that is costly to acquire, transfer and use elsewhere 'sticky'. Lippman and Rumelt showed that when the causes of a firm's performance are ambiguous, imitation stays uncertain.

What AI rebuilds show. The gap is visible in public results. The compiler of version 0.1 cost under $20,000 and still emits slower code than GCC with optimisation off. On GSO, 102 optimisation tasks from 10 codebases, leading agents matched the expert's result in fewer than 5% of attempts when it was published in 2025. On PaperBench, the best agent scored 21% at replicating 20 research papers from scratch. Function arrives first. Quality arrives later, and at a cost.

Quality parity. A replica reaches quality parity when it matches the original on a declared public measure within a stated tolerance: accuracy against public ground truth, speed, error rate, forecast skill. Every ORC figure now names its quality target as well as its fidelity level.

Decision points. Finesse is sized the way function is. Function is counted in function points. Finesse is counted in decision points: choices in the quality-critical logic that move the quality measure and cannot be read from public sources. Constants, thresholds, weights, rule-table rows, ordering rules and special cases all count.

Class	Test	What it costs a replicator	As AI improves
Derivable	Documented, or standard practice	Nothing beyond the function itself	No change
Discoverable	Changing it changes what an outside observer can see	A search, priced per decision point by the Index	Falls with token prices
Hidden	Changing it changes quality, but not what an outsider can observe in time	Independent re-creation, then proof of parity	Re-creation falls. Proof takes calendar time and does not

The cost. Four terms are added to the build B of version 0.1.

$$B_\varphi \;=\; Z_h\,u \;+\; N_d\,s_O \;+\; N_h\,s_R \;+\; L\,(p + a_v\,w), \qquad \varphi \;=\; \frac{B + B_\varphi}{B}$$

Symbol	Meaning	Source
Z_h	Hidden work: function points the public manifest missed	Sealed measurement; a declared range without one
N_d, N_h	Discoverable and hidden decision points	Sealed measurement
s_O	Cost to find one decision point at observability grade O	The Index
s_R	Cost to re-create one hidden decision point to the quality target	The Index
L	Verification lag: months needed to prove parity	Computed from the quality measure
a_v	Attention while verifying, hours a month	The Index
φ	Finesse factor: total build over the function-only build	Derived

Proof time. The lag is a statistical fact about the quality measure, not an opinion.

$$L \;=\; \Delta\left(\frac{(z_\alpha + z_\beta)\,\sigma}{\varepsilon}\right)^2$$

Δ is the time between independent public observations of quality, σ their noise and ε the tolerance. With weekly observations and noise three times the tolerance, proof at 5% significance and 80% power needs 56 weeks, about 13 months. During the lag a replicator pays for a seat on the original and for the attention to compare.

Finesse grades. Until the Index measures finesse, a public assessment assigns a grade and a declared factor.

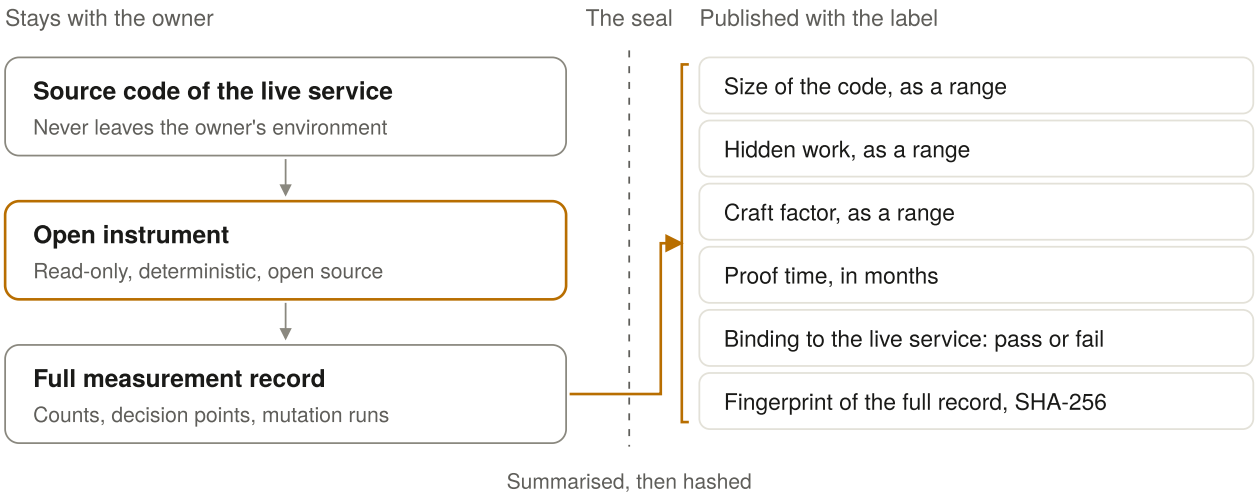
Grade	Name	Evidence	Declared factor
Φ1	Plain	Standard methods, documented behaviour	1.0 to 1.1
Φ2	Tuned	Visible calibration or corrections that the outputs reveal	1.1 to 1.5
Φ3	Crafted	Many interacting undocumented rules	1.5 to 3
Φ4	Expert	A measured quality lead over the best open substitute, resting partly on hidden decisions	Above 3, plus a verification lag
Φ5	Singular	No public quality measure, or proof takes longer than the horizon	Not priced. Reported as the sealed share

The rule that follows. Discovery gets cheaper with AI. Proof does not. Finesse an outsider can observe is rediscovered ever more cheaply, so it lengthens the runway by months. Finesse that can only be proven over time puts a floor under the price.

Hidden finesse counts only against a public quality measure. A claim of excellence that nobody can check adds nothing to the score.

Open method, sealed evidence

The method, the instrument and the Index are public. The owner's code never is. A sealed measurement lets the real code inform the score while only inferences leave the owner's environment.



Sealed measurement: what stays with the owner, and what is published.

At level A4 the run takes place in front of an accredited assessor, or inside an attested enclave.

Why a seal. Without it the public tier has to guess hidden work and finesse, and the range stays wide. With it the owner trades evidence for a narrower range. Other fields already work this way: the evidence stays private and the conclusion is published.

Field	What stays private	What is published
Security certification	Source code, reviewed by an accredited laboratory at the higher assurance levels	The certification report, the certificate and the security target, on a public portal
Service audits	The SOC 2 report: controls tested and results, restricted to named parties	The SOC 3 report: the auditor's opinion, for general use
Health technology appraisal	Commercial-in-confidence data, redacted	The guidance and its reasoning. Data fundamental to the decision cannot be marked confidential
Confidential computing	Code and data inside an enclave	A hardware-signed attestation of which public checker ran on them

Four assurance levels. Each label states how much evidence stands behind it.

Level	Name	Evidence	Who can check it	Effect on the range
A1	Public	The outside view only	Anyone, from open inputs	Widest: hidden work and finesse are declared ranges
A2	Declared	The owner's signed statement: manifest corrections, data classes, time-locked assets, the quality measure	Anyone can read it. Nobody has verified it	Corrects scope. Does not narrow cost
A3	Sealed	The owner runs the open instrument on the real code and keeps the record. Its hash is published	An arbiter, in a dispute	Measured counts replace the declared ranges
A4	Attested	An accredited assessor witnesses the run, or it runs inside an attested enclave	The assessor signs. The Keeper re-runs a sample	As A3, with independent assurance

The instrument. It is open source, runs read-only inside the owner's environment, and does seven things.

1. Hidden work. It counts function points in the code and compares them with the public manifest.
2. Decision points. It lists constants, thresholds, weights, rule rows and special cases in the quality-critical code.
3. Materiality. It perturbs each one and re-runs the quality harness. Only those that move the public quality measure beyond tolerance count.
4. Identifiability. It checks whether the perturbation changes what an outsider can observe. If so the decision is discoverable. If not it is hidden.
5. Interdependence. It samples pairs to see whether one decision's best value depends on another, which makes a search harder.
6. Proof time. It computes the verification lag from the noise of the public quality measure.
7. Binding. It confirms that the measured code reproduces the live service's public behaviour on a random probe set.

Steps 3 and 4 are mutation testing pointed at a new question. They need no judgement, so two assessors get the same counts from the same code.

What leaves, and what stays.

Tier	Contents	Who sees it
Public label	P/R with ranges, Value Score, finesse grade, sealed share as a band, assurance level, version stamps, hash of the sealed record, assessor and funding	Everyone

Tier	Contents	Who sees it
Owner's report	Where the cost sits by capability, counts by class, the discoverable share, exposure thresholds, the runway, the durable price	The owner
Sealed record	Raw instrument output, bound to the commit and to the harness	Kept by the owner. Opened only to an arbiter
Source code	Not copied, not transmitted	Nobody outside the owner's environment

Rules that keep the seal honest.

- No judgement under seal. The instrument is deterministic: same code, same version, same counts.
- A public anchor. Sealed finesse counts only against a quality measure that outsiders can check.
- Consistency. A sealed median outside the public range goes to a second, attested run before it is published.
- Bands, not counts. Public outputs are too coarse to serve as a build plan.
- Commitment. The hash fixes the record on the day of assessment. A dispute opens that record, not a new one.
- Sanction. A false binding or a tampered harness withdraws every score of that owner, publicly.

The error range

Every figure carries an 80% range, and the range says where its uncertainty comes from.

How it is computed. The checker draws each uncertain input from a declared distribution and recomputes the result 100,000 times with a fixed seed. It reports the median with the 10th and 90th percentiles. Anyone who runs the same inputs gets the same numbers.

Where it comes from. The worked example later in this document splits its range as follows.

Source of uncertainty	What settles it	Share of the range today	Share at month 36
Unit costs: tokens and oversight per point	A measured Index	60%	11%
Attention needed to run a replica	A measured Index	18%	34%
Size count	A second count, or a sealed count	10%	2%
Hidden work	A sealed measurement	7%	1%
Finesse	A sealed measurement	6%	1%

Source of uncertainty	What settles it	Share of the range today	Share at month 36
Pace of AI progress	Time	0%	51%

Shares of the variance of the logarithm of replication price, each input varied alone. They are rounded, so a column may not sum to 100.

Today most of the range is the missing Index. Three years out, half of it is the pace of AI, which no measurement removes.

Three rules.

- A figure quoted without its range is not an ORC figure.
- Declared inputs widen the range. They never move the median in the owner's favour.
- A sealed measurement replaces a declared range with a measured one. It can raise the median or lower it.

The range is itself tested. When a scored product is later rebuilt in public with disclosed costs, the register records where the outcome fell. An 80% range should contain eight outcomes in ten. The Keeper publishes that hit rate every year, so the score earns trust the way a forecaster does: by a record.

The Value Score

The Value Score is the discount to replication price that a subscription still offers 36 months from now, at the cautious end of the range. It is issued only when that discount is positive in at least nine outcomes out of ten.

$$\text{Value Score} \;=\; 100\left(1-\frac{S}{\text{ORC}_{36}^{P10}}\right), \qquad \text{issued only if } \Pr[\,S<\text{ORC}_{36}\,]\geq 0.90$$

ORC36 is the replication price of a build that starts 36 months from now, on the Index trend. P10 is its 10th percentile. A score of 40 says: after three more years of AI progress, and even if replication turns out as cheap as the range allows, subscribing still costs 40% less.

Why this test. A buyer can switch to a replica in any month, so the long-term question is whether some future build date beats staying subscribed. For a need that continues, that happens when forward P/R passes 1. The month it does is the runway of version 0.1. The score asks whether the runway is longer than the horizon, with 90% confidence.

Four results. The advice still runs in one direction only.

Result	Condition	What the label says
Premium	Spot P/R above 1	The premium and what it buys. Never advice to clone
Subscribe	Spot P/R below 1, long-term test not passed	P/R today, and the month until which the value holds at 90% confidence
Value Score	Passes at month 36 with 90% confidence	The score, from 1 to 100
Value Score, durable	Also passes at the floor with 90% confidence	The score, marked durable: tokens could be free and the plan would still be cheaper

What every label prints. Replication price with its range. P/R at the three dates. The hold probability: the share of simulated outcomes in which subscribing is still cheaper at 12, 24 and 36 months. The month the value holds until. The finesse grade, the sealed share, the assurance level and the version stamps. Then the Value Score, or the plain statement that none is issued.

What the score leaves out. It counts only what can be priced. Content, trust and anything the seal marks as not replicable appear beside it as the sealed share and are never folded in. That keeps the number conservative, and it means that adding one unreplicable feature does not lift it.

Three ways to earn it.

1. Price at or below the durable price: the highest price the floor supports.
2. Prove under seal a component whose quality takes calendar time to verify, against a public measure.
3. Supply something a replica must keep paying for: licensed data, live operations, continuous access to the original.

$$p_{\text{durable}} = \frac{\text{ORC}_{\infty}}{k\, A_T}$$

The gate is strict on purpose. P/R describes every plan. The Value Score is a seal for the plans whose value survives the pace of AI.

Worked example: COTBase Premium under version 0.2

On public evidence COTBase Premium costs about a tenth of its replication price today and keeps that advantage for 27 months at 90% confidence. It does not yet earn a Value Score.

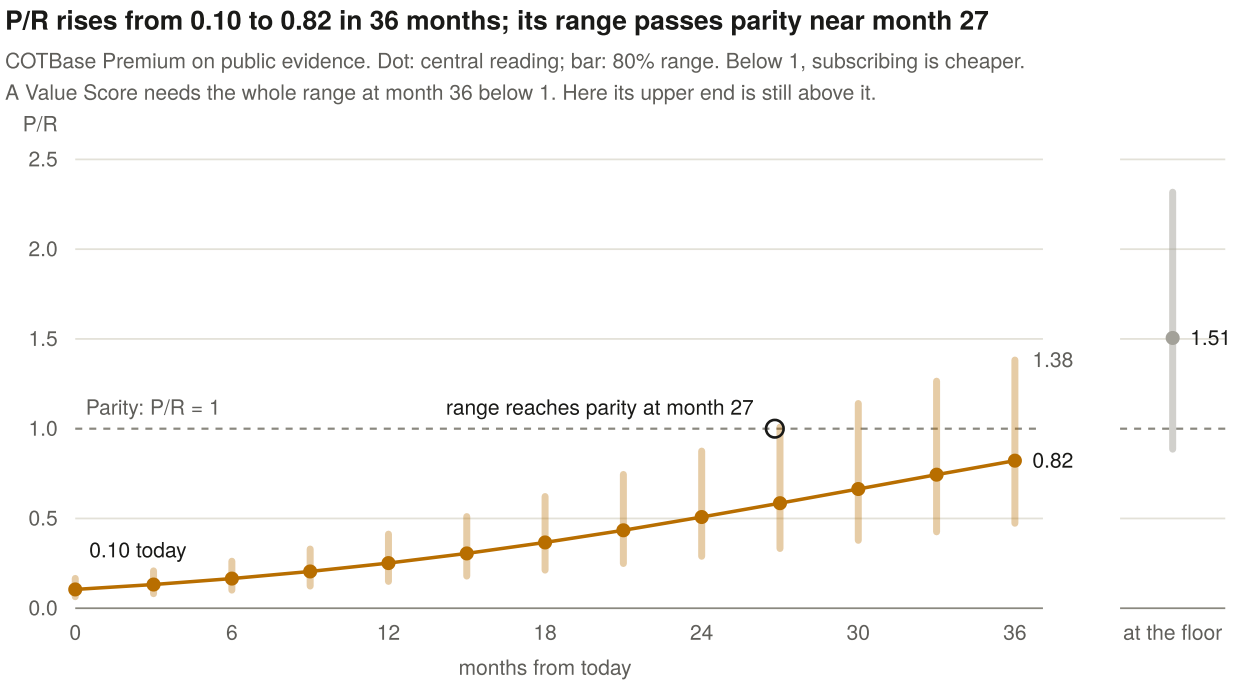
An illustration, not a certified score. The assessment is at level A1. Nobody has inspected the COTBase code. Unit costs, hidden work and finesse are declared ranges, and every result is provisional.

Inputs. The low and high cases of version 0.1 become the 10th and 90th percentiles. Three inputs are new.

Input	10th to 90th percentile	Grade
Price; wage	$14 a month on the yearly plan; $65.38 an hour	A
Size of the standard buyer's core	80 to 125 function points	C
Tokens and oversight per point	$2 to $6 and 0.08 to 0.24 hours, moving together	C
Hidden work (new)	5% to 40% of the counted size	C
Finesse factor (new)	1.1 to 1.5: grade Φ2, for the corrected and combined series that the outputs reveal	C
Attention to run a personal copy	0.25 to 1 hour a month today; 0.05 to 0.2 at full automation	C
Halving time of AI-priced cost (new as a range)	6 to 12 months	C

The label.

Field	Result on public evidence
Replication price today	$4,800, range $3,000 to $7,900
Hold probability at 12, 24 and 36 months	100%, 94%, 68%
Value holds until	27 months at 90% confidence; 43 months at the median. Parity never comes in 16% of outcomes
Token share of today's replication price	19%, range 13% to 23%
Durable price	About $9 a month at the median floor; under $6 at the cautious end
Finesse grade and sealed share	Φ2, declared. No sealed share evidenced
Assurance level	A1, public
Value Score	Not issued. The plan holds at month 36 in 68% of outcomes, and the test requires 90%



P/R of COTBase Premium over the next 36 months, from orc_value_model.py with seed 20261113 and 100,000 draws, on public evidence (level A1), for an outlay of $504 over 36 months.

P/R climbs towards parity because the part of the cost that AI prices keeps halving. The grey mark is the terminal reading: the same outlay against the floor that remains once that part has gone.

What the result says.

1. Subscribe. A personal copy costs about ten times the plan today.
2. The advantage is real but not permanent. Most of today's hurdle is priced in tokens and oversight, and both are falling.
3. What remains at the floor is the user's own attention: an hour to commission a copy and a few minutes a month to mind it. That supports about $9 a month, against a price of $14.
4. Tokens are still the small part, at about a fifth of the cost.

What a sealed measurement could change. The findings below are invented to show the mechanics.

Scenario	Replication price today	Value holds until, at 90% confidence	Hold probability at month 36	Value Score
A. Plain: 8% hidden work, finesse 1.05	$4,000, range $2,700 to $6,200	25 months	62%	Not issued
B. Crafted: 30% hidden work, finesse 1.8	$6,600, range $4,300 to $10,600	30 months	78%	Not issued

Scenario	Replication price today	Value holds until, at 90% confidence	Hold probability at month 36	Value Score
C. As B, plus a hidden component whose quality takes 13 months to prove	$7,300, range $4,900 to $11,200	Beyond the horizon	Over 99%	26
B, with a measured Index	$6,500, range $5,400 to $7,800	37 months	94%	4
C, with a measured Index	$7,100, range $6,000 to $8,500	Beyond the horizon	100%	38, durable

- A seal can cut both ways. Scenario A lowers the owner's number.
- Visible finesse buys months. Scenario B adds three months to the 90% runway and earns no score.
- Proof time changes the result. In scenario C the floor rises from about $335 to about $620, above the $504 outlay, and a score is issued.
- Measurement compounds. A measured Index cuts the spread of today's price from 2.6 times to 1.7 times. A sealed measurement on top cuts it to under 1.5 times, and scenario B then passes narrowly.

The rest of the product. The group and entrant tests of version 0.1 are unchanged: a shared living replica matches the price at 35 to 58 members. iCOT stays outside the score because its algorithm is hidden. If each week's published report can serve as a public check on an intraday estimate, iCOT has a quality measure and a verification lag, which is the profile of scenario C. That is an inference from the public description, not a finding.

The upkeep principle: built value expires

Anything built is a decaying asset. The price of rebuilding it falls every month, so a price that is justified today stops being justified unless its creator keeps renewing the reason. This section is proposed for version 0.3.

The precedent. Value has been made to expire on purpose before. In Wörgl in 1932 a local note stayed valid only if its holder bought a stamp for it each month, at 1% of its value. In post-war Germany a food ration stamp was valid for one month. A built service is a note of the first kind. It keeps its value only while someone keeps paying for the stamps.

The law of decay. The replication price has two parts: the floor, which AI does not make cheaper, and the hurdle above it, which AI prices. The hurdle halves every h months, whether or not anyone copies the service.

$$A(t) = \mathrm{ORC}(t) - \mathrm{ORC}_{\infty}, \qquad \frac{dA}{dt} = -\lambda A, \qquad \lambda = \frac{\ln 2}{h}$$

Halving time	Lost each month	Lost each year	To stand still, rebuild the whole hurdle every
6 months	10.9%	75%	8.7 months
8.5 months, the declared median	7.8%	62%	12.2 months
12 months	5.6%	50%	17.3 months

The score, dated. The PayOrBuild score is the headroom above the line, at the cautious end of the range, for a build that starts in month t.

$$\Sigma(t) \;=\; 100\left(1-\frac{S}{\mathrm{ORC}_t^{P10}}\right)$$

Above zero the service is above the line: subscribing is cheaper in at least nine outcomes out of ten. The Value Score is this score read 36 months ahead. A score is always quoted with its date, because it is falling.

The upkeep condition. A service stays above the line only if what it adds each month makes up for what decay removes. Three things count: new replication value m, a rising floor and a falling price.

$$m \;+\; \frac{d\,\mathrm{ORC}_\infty}{dt} \;-\; T\,\frac{dp}{dt} \;\;\geq\;\; \lambda\left(S-\mathrm{ORC}_\infty\right)$$

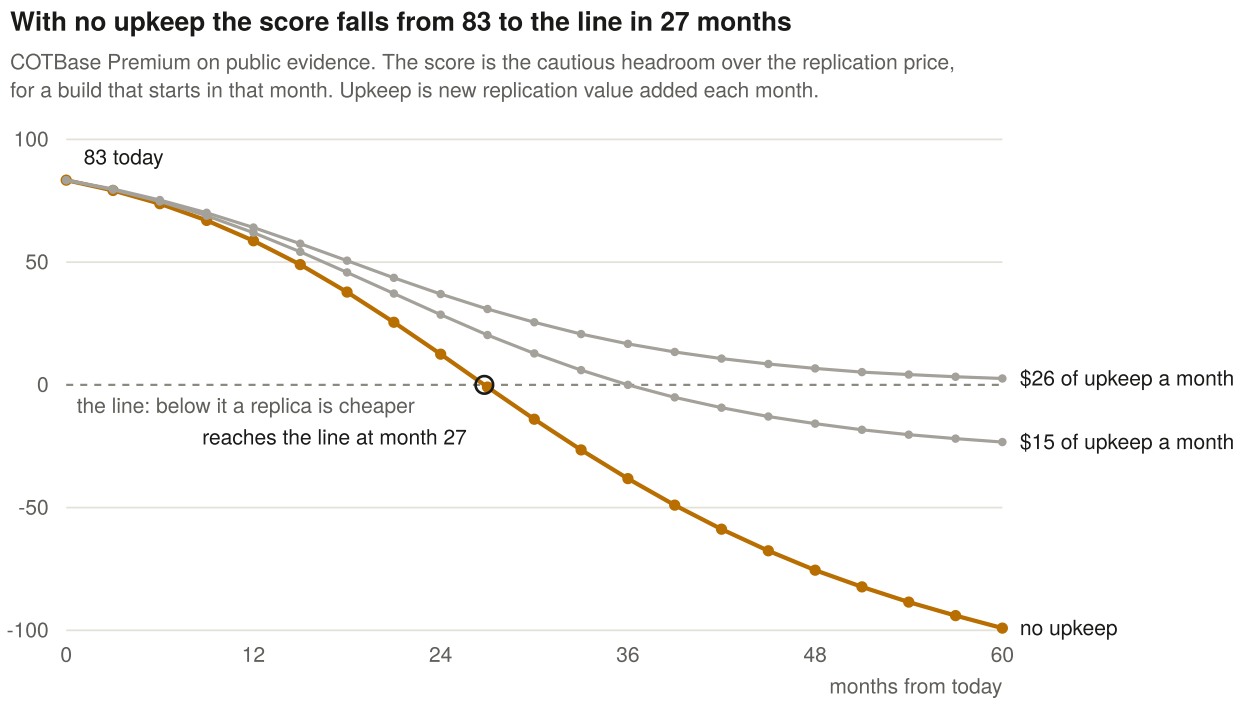
The right-hand side is the stamp: the rate of decay times the part of the outlay that the floor does not cover. A service whose floor is above its outlay owes no stamp.

Three ways to pay the stamp.

Way	What the creator does	What it costs	How long it works
Run	Adds capability that subscribers use	The same dollars each month, but twice the function points every halving time	A few halving times, until the points needed outgrow what a buyer can use
Anchor	Raises the floor: proof that takes calendar time, licensed data, live operations	A running cost that every replica must also pay	For good
Follow	Lowers the price along the line	Revenue	For good, down to the durable price

The creator's advantage. The creator pays for an addition once. Every replicator would have to pay for it again. That makes the stamp cheap for a service with many subscribers. It works only for additions the standard buyer uses, because capability nobody uses does not count towards size.

COTBase Premium with no upkeep. On public evidence the score falls from 83 today to the line at month 27.



The score of COTBase Premium with no upkeep, from upkeep.py on the same model and seed, on public evidence (level A1).

Reading	Today	Month 12	Month 24	Month 36
The line, as a monthly price	$84	$34	$16	$10

The line as a price is the highest monthly price at which the plan stays above the line. It falls towards $6, the cautious floor. The plan's $14 is under it until month 27.

What the stamp costs here.

Upkeep	New replication value a month	The same in function points
The stamp at the line, at the median	$13	0.4 a month today, 8 a month by month 36
Keeps the score above the line at month 36	$15	0.5 a month today, 10 by month 36
Keeps it above the line for good	$26	0.9 a month today, 17 by month 36, 44 by month 48
Holds today's score of 83	$362	12 a month today, doubling every 8.5 months

Upkeep	New replication value a month	The same in function points
An anchor instead: a running cost every replica must also pay	$8	None

In dollars the stamp is small: one to two months of the plan's price in new replication value each month. In function points it doubles every halving time, and by the fourth year it asks for close to half of today's core every month. Running buys time. Anchoring and following are the upkeep that lasts.

What the label adds.

- Valid until: the month in which the score reaches the line with no upkeep.
- The stamp: the upkeep a month that holds the line.
- The line as a price, today and 36 months ahead.

A score already lapses after 90 days, so a creator who displays the mark renews it four times a year.

The rule. A service that asks for money should keep its score above the line. The standard does not enforce that. It dates every score and prints the stamp, and buyers and their agents do the rest.

Figures in this section come from upkeep.py, a short script on the same model and seed.

The business model

The standard gives away what tokens can copy and sells what they cannot. The method, the instrument and the Index are free. The paid products are independence, a signature, a live register and a track record.

The standard applied to itself. Its own function and data are layers 1 and 2, which the framework says fall towards zero. So nothing is charged for them. Revenue comes from layers 3 to 6: operations, the stream of updates, judgement and trust. Each year the Keeper publishes the P/R of its own paid products.

Five roles.

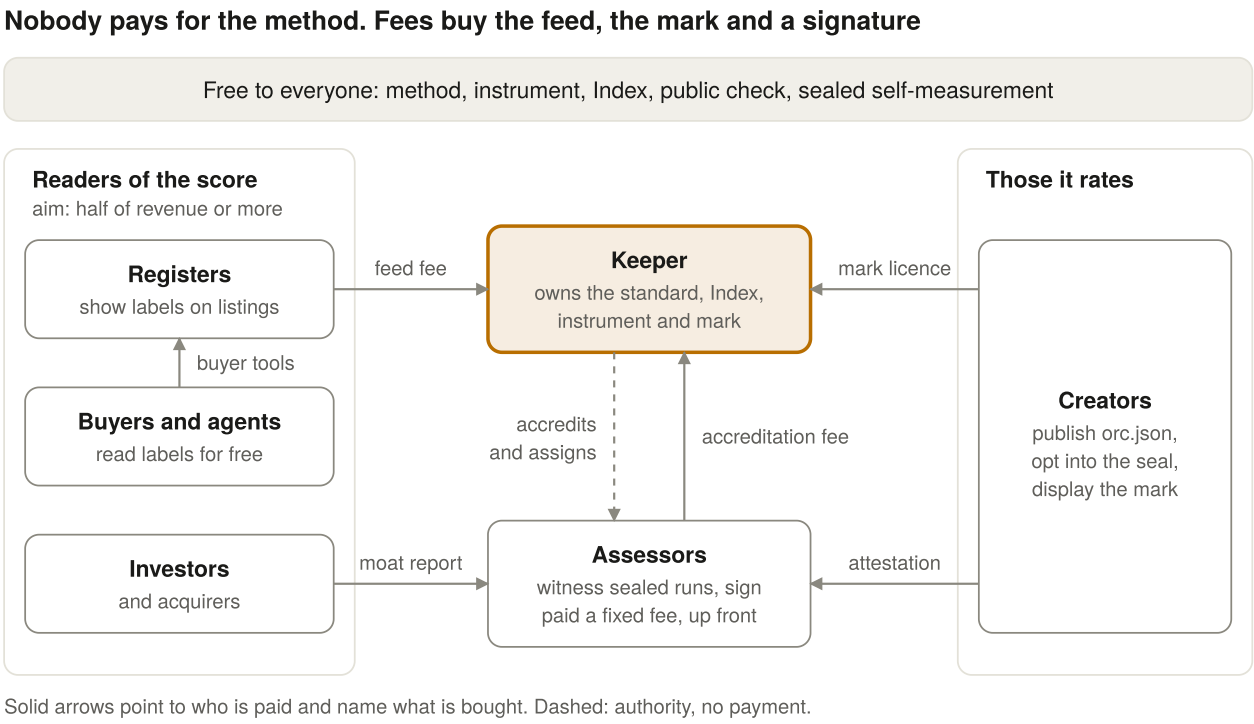
Role	What it does	Paid by	Must not
Keeper	Owns the standard, the Index, the instrument and the mark. Accredits assessors. Publishes calibration. Runs disputes	Mark licences, feed fees, accreditation fees, grants	Assess products, or take a fee that depends on a result
Assessor	Runs public checks. Witnesses sealed measurements. Writes the owner's report	Whoever commissions the work: owner, buyer, register or investor	Pick its own subjects when the owner pays, or hold code

Role	What it does	Paid by	Must not
Register	Shows labels beside products: marketplaces, comparison sites, app stores, procurement tools	Its own business	Let ranking depend on commission
Creator	Publishes orc.json. Opts into a sealed measurement. Displays the mark	Its customers	Alter the instrument or the quality harness
Buyer, and the buyer's agent	Reads labels. Runs its own check	Nobody	Nothing: reading is free

Nine products.

Product	Buyer	Price logic	What is sold
Public check (A1) and declaration (A2)	Anyone	Free. The tool is open, and the cost is tokens	Nothing: function is free
Sealed self-measurement (A3)	Creator	Free. The owner runs the open instrument	Nothing
Attestation (A4)	Creator	A fixed fee by size band, set by the Keeper and paid before the run	Independence and a signature
Mark licence	Creator with a Value Score	Yearly, by plan revenue band. Free below a threshold	Trust
Register feed	Marketplaces and comparison sites	A flat fee by listing count. Free for non-commercial use	The stream, and operations
Buyer tools	Procurement teams and buyers' agents	A subscription: portfolio P/R, renewal alerts, parity briefs for negotiation	The stream, and judgement
Moat report	Investors and acquirers	Per engagement, under seal: the share of revenue on plans above forward parity	Judgement and trust
Parity licence exchange	Creators and large buyers	A success fee on a source or self-host licence priced under ORC	Network
Accreditation and training	Assessors	A fee	Trust

Who sells what. The Keeper sells the mark licence, the register feed and accreditation. Assessors sell attestation and moat reports. Registers sell buyer tools and run the licence exchange inside their own business.



The business model: who pays whom, and for what. The licence exchange is not drawn.

Four pricing rules.

1. Never sell the direction of a number. Fees are fixed and paid before the result exists.
2. The Keeper assigns the assessor. The owner does not choose.
3. Aim for at least half of revenue from readers of the score, not from those it rates.
4. Publish the standard's own P/R every year.

Why these rules. Credit rating agencies began in 1909 on a subscriber-pays model and moved to issuer-pays in the 1970s, which an SEC commissioner describes as a fundamental conflict of interest. After the financial crisis the SEC was asked to study assigning the rater instead of letting the issuer choose. Consumer Reports has rated products since 1936, accepts no advertising and pays for what it tests. Rules 1 to 3 take the second model and guard against the first.

Why the conflict is smaller here. A rating is an opinion. An ORC assessor witnesses a deterministic measurement that any other assessor would reproduce. Shopping for a kinder assessor changes nothing.

What it costs to run. A public check sizes and prices a product without building anything, so it costs tokens and minutes. An Index reference run costs about what the replication it measures costs: $200 to $1,800 in tokens for a service of 100 to 300 function points, at the declared $2 to $6 a point.

The launch needs no new company. Phase 1 is a feature of businesses that already exist: one consenting product as the first subject and one marketplace as the first register. Where the register's owner also owns a listed product, that product's label says so, shows only A1 or A3 results with every input, and carries no mark until an independent assessor exists.

Incentives, conflicts and gaming

The design works only if the people it measures gain by telling the truth. Four incentives carry it, and every new feature brings an attack to block.

Four incentives.

Who	Incentive	Why it holds
Creators with real finesse	Opt into the seal	When disclosure can be verified, good news is disclosed and silence is read as bad news. This is the unravelling result that Dranove and Jin review. The sealed owner lifts its cautious bound; the rest stay at the public estimate
All creators	Price lower, not higher	The Value Score rises as P/R falls, so competing on the score pushes prices down
Assessors	Compete on cost and speed, not leniency	Results are reproducible, so a kinder assessor has nothing to offer
Registers	Sort by the label	It is an order that sellers cannot buy

The same review warns that certifiers do not always report accurately or without bias. The rules below assume they will be tempted.

Ten new attacks. These add to the nine of version 0.1.

Attack	By	Safeguard
Finesse padding: fill the code with constants to inflate decision points	Creator	Materiality: only decisions that move the public quality measure count
Harness tuning: a private quality test that flatters the code	Creator	The quality measure must be checkable from outside. The harness hash is sealed with the record
Wrong code: measure a showcase repository	Creator	Behavioural binding to the live service on a random probe set
Manufactured opacity: hide outputs so that discoverable decisions look hidden	Creator	Hidden finesse needs a measured quality lead over the best open substitute. Friction is reported apart, as in version 0.1

Attack	By	Safeguard
Selective sealing: publish only favourable results	Creator	A sealed run commits to publication before the result is known
Leakage: read the public bands as a build plan	Replicator	Bands, minimum aggregation, no output below the level of the plan
Assessor capture	Assessor and creator	Assignment by the Keeper, fixed fees, sampled re-runs, a deterministic instrument
Price signalling: under-priced creators read the label as room to raise prices	Creators	The label is a cost benchmark from public inputs, never a price recommendation, and the score falls as the price rises
Register bias: rank by commission	Register	A published ranking rule, disclosed commissions, feed withdrawn on breach
Trend lobbying: argue for a slower pace of AI	Creators	The trend is computed from the Index by a published formula

Becoming the standard unit

Units of comparison win in three ways: a law requires them, a dominant intermediary adopts them, or a free tool makes them the default. The third path is open now, and buyers' agents are its engine.

How other units won.

Unit	How it became standard	Lesson for P/R
Annual percentage rate	Required of US lenders by the Truth in Lending Act, 1968	Law follows a unit that already works
Unit price	Required beside the selling price by EU Directive 98/6/EC	Comparison needs one denominator
Michelin stars	A tyre maker's free guide for motorists from 1900. Stars from 1926, awarded by anonymous inspectors	A company with an adjacent interest can found a standard if inspection is independent
Financial benchmarks	Rebuilt on the IOSCO principles of 2013 after attempted manipulation of major rates	Governance, method and accountability must be public before trust returns

Why agents decide it. A person rarely computes build-or-buy. An agent does it every time it is asked to renew a subscription. It needs a number it can read and recompute, and it will prefer vendors that publish one. The orc.json statement of version 0.1 is that number's address.

The adoption loop.

1. The free checker scores products at level A1.
2. Registers show P/R and sort by it.
3. Owners with real finesse seal their code to narrow the range.
4. Sealed scores build the calibration record.
5. The record makes the unit safe to write into procurement rules and contracts.

Four phases.

Phase	What happens	Exit test
1. Proof	Checker and instrument published as open source. COTBase scored at A1 and, with its owners' consent, measured under seal. One marketplace shows labels	Two independent recounts of one manifest agree within tolerance
2. Register	20 to 50 consenting products in one category. Index epoch 1. orc.json published by each	A buyer or an agent outside the founders' group uses a label
3. Independence	A neutral Keeper. A first accredited assessor outside the founders. Mark licences begin	The first A4 attestation of a product the founders do not own
4. Standard	Version 1.0. Buyer tools and moat reports. Calibration published yearly	The hit rate of the 80% range published on at least 30 outcomes

The counts in phases 2 and 4 are proposals, not forecasts.

Governance, legal exposure and limits

The standard's weak points are its founders' conflicts, the law on comparative statements, and the honesty of a three-year forecast.

Governance rules added in version 0.2.

Topic	Rule
Keeper	Independent of every assessor, register and rated product. The founders hold no majority
Founders' own products	Labelled as such. No mark until an assessor outside the founders attests them
The Index as a benchmark	Run on the pattern of the IOSCO principles: published method, hierarchy of inputs, change procedure, conflicts policy, audit

Topic	Rule
Trend rule	The halving time is computed from the last four Index epochs by a published formula
Sealed records	Kept by the owner under a published hash for 24 months. Opened only to an arbiter
Assessors	Accredited, assigned, rotated, and sampled for re-runs. Liable for a false binding
Stamps	Standard version, Index epoch, instrument version, manifest version and date. Scores lapse after 90 days

Legal exposure. This is a map, not legal advice, and it needs a lawyer before any label names a product without its owner's consent.

Issue	Why it arises	Design answer
Comparative statements	In the EU a comparison that identifies a competitor's product is permitted only if it is not misleading, compares like with like, and objectively compares verifiable features, which may include price	Open inputs, recomputable results and standard conditions. A register owned by a competitor of a listed product needs particular care
Price coordination	Pricing software built on competitors' non-public data has drawn antitrust action in the US	Public inputs only. No price recommendations. Owners' reports are never shared or pooled
Trade secrets	A seal that leaks is worse than no seal	The code never leaves. Bands only. Assessor liability
A wrong score	A poor label can cost sales	Ranges on every figure, a notice period, a dispute procedure, a published method

Limits.

- The forward test is a forecast. A halving time of 6 to 12 months is declared, not measured, and the hold probability depends on it more than on anything else.
- Decision points are a new measure. Two instruments may count differently until a reference suite fixes the rules.
- The seal needs a quality harness. A product with no measurable quality cannot use it.
- The floor rests on what a user's attention is worth. For a buyer who values that time at zero the floor is lower.
- Hidden finesse can be re-invented. An AI that independently builds something as good is a competitor, not a replicator, and no score protects against it.
- A strict gate will issue few scores. That is the price of a seal that means something. P/R still describes every plan.
- The founders are judge and party until the Keeper is independent.

Open questions.

1. Should the forward date be 36 months in every category, or follow each category's pace of change?
2. What tolerance defines quality parity, and who sets it for each category?
3. How coarse must the public bands be to stop leakage?
4. Should a personal label use the buyer's own wage in place of the reference wage?

Roadmap

Six steps take version 0.2 to a first attested score, and the first three need only what the founders already have.

Step	Work	Result
1	Publish the value model beside the version 0.1 model. Add P/R, hold probability and runway to the checker	Labels at level A1
2	Write the instrument: function-point and decision-point counters, materiality and identifiability runs, the binding test	An open tool an owner can run alone
3	Run it on COTBase against a declared quality measure. Publish the label and the hash	The first A3 score, and the first real finesse count
4	Show labels on one marketplace. Publish the ranking rule and the commission disclosure	The first register
5	Index epoch 1, with finesse curves: the cost of lifting a working replica to a given share of the original's quality, on open references	Measured unit costs replace the declared ranges
6	Constitute the Keeper. Accredit an assessor outside the founders. Publish the first calibration record	Independence, and the first attested Value Score

Sources

Every figure in the worked example comes from the value model, a short Python script delivered with this document as orc_value_model.py. It runs with a fixed seed, so the same inputs give the same numbers. The theory it extends is in PayOrBuild.

Units and standards

- Truth in Lending, Comptroller's Handbook, US Office of the Comptroller of the Currency
- Directive 98/6/EC on the indication of the prices of products offered to consumers
- History of the MICHELIN Guide, Michelin
- Principles for Financial Benchmarks, IOSCO, July 2013

Private evidence, public result

- Common Criteria security evaluation, Secura factsheet
- SOC 2 and SOC 3 reports: what is the difference?, Linford and Co
- Guide to the processes of technology appraisal, section 3, NICE
- Confidential Attestation: Efficient in-Enclave Verification of Privacy Policy Compliance, arXiv

Certification, disclosure and conflicts of interest

- Quality Disclosure and Certification: Theory and Practice, Dranove and Jin, Journal of Economic Literature, 2010
- Statement on the removal of references to credit ratings from Regulation M, SEC Commissioner Crenshaw, June 2023
- Agenda for the Credit Ratings Roundtable, SEC, April 2013
- Consumers Union acquires The Consumerist, Consumer Reports, for its funding policy

Imitation and tacit knowledge

- Capturing Value from Technological Innovation, Teece, a later version of his 1986 paper in Research Policy
- 'Sticky Information' and the Locus of Problem Solving, von Hippel, Management Science, 1994
- Uncertain Imitability, Lippman and Rumelt, Bell Journal of Economics, 1982

The quality gap in AI rebuilds

- GSO: Challenging Software Optimization Tasks for Evaluating SWE-Agents, NeurIPS 2025
- PaperBench, OpenAI, April 2025
- Building a C compiler with a team of parallel Claudes, Anthropic

Law

- Directive 2006/114/EC, Article 4, on misleading and comparative advertising
- DOJ's RealPage settlement: a blueprint for 'safer' algorithmic pricing?, Fenwick

Expiring value

- Good ideas that didn't pan out: free money, The Long and Short, on the stamp scrip of Wörgl
- The Wörgl experiment in Tyrol, Cité de l'Économie
- Lebensmittelmarken, transcript, EDU Westfalen, on post-war ration stamps valid for one month

---

The ORC Value Standard, version 0.2, published by PayOrBuild on 13 November 2026, for comment. Every figure in it is an estimate with a stated range.