# Package: tokenusagedata (via r-universe)

November 19, 2026

**Title** AI Token Prices, Context Windows and Model Sizes

**Version** 0.1.0

**Description** Provides five documented data frames for analysing artificial intelligence (AI) model application programming interface (API) prices, published context windows, and published or estimated parameter counts. Fixed October 2026 snapshots retain source URLs and qualifications and add typed dates and numeric quantities for reproducible comparisons and teaching. Data are from Pistachio Technology Inc. (2026), published at <https://tokenusage.dev/publications>: API prices <doi:10.5281/zenodo.23552904>, context windows <doi:10.5281/zenodo.23552975>, and model sizes <doi:10.5281/zenodo.23553072>. These are selected historical records; estimated sizes are distinguished from lab disclosures.

**License** CC BY 4.0

**Encoding** UTF-8

**Depends** R (>= 3.5)

**LazyData** true

**LazyDataCompression** xz

**NeedsCompilation** no

**URL** https://tokenusage.dev/publications

**Author** Jack He [aut, cre], Pistachio Technology Inc. [cph]

**Maintainer** Jack He <jack.beanstalk.2022@gmail.com>

**Repository** https://cran.r-universe.dev

**Date/Publication** 2026-11-19 15:00:02 UTC

**RemoteUrl** https://github.com/cran/tokenusagedata

**RemoteRef** HEAD

**RemoteSha** 8f06207dd42643028b3a2c6138b362f0c8ed27c0

# Contents

tokenusagedata-package . . . . . 2
context_windows . . . . . 3
model_size_estimates . . . . . 5
model_sizes . . . . . 6
token_prices . . . . . 8
token_prices_premium . . . . . 10

**Index** **12**

---

`tokenusagedata-package`

*AI Token Prices, Context Windows and Model Sizes*

---

## Description

Five documented historical data frames from tokenusage.dev research notes 002, 003 and 004, with source text, qualifications, typed dates and nominal numeric quantities.

## Details

The package contains `token_prices`, `token_prices_premium`, `context_windows`, `model_sizes` and `model_size_estimates`. They derive from five CSV files in three fixed Zenodo version records. The October 2026 snapshots are selected historical records, not a complete market census or a live service.

Data are available offline. All tables are base R data frames with automatic row names. Empty CSV fields become missing character values. Original text retains its UTF-8 characters.

Dates use a leading YYYY-MM-DD when present; partial dates remain missing, and date qualifications are retained in the tables that have qualified date fields. Nominal context and parameter counts come from the first figure in the corresponding text label. Commas and an optional leading tilde are removed; k, M, B and T denote powers 3, 6, 9 and 12 of ten. This extraction does not reconcile conflicts, measure effective context, or establish the accuracy of an estimate.

Prices are US dollars per one million tokens. Published unrounded blends weight input and output 3:1. Context units and access conditions differ across models. Published parameter counts and attributed closed-model estimates remain separate. Consult each table's help before making comparisons.

Only the five identified CSV files are included. Excluded source rows, other downloads, article text, charts and images are not package data.

The compilation is credited to tokenusage.dev and was deposited by Pistachio Technology Inc. The data and this package are licensed under Creative Commons Attribution 4.0 International (CC BY 4.0): https://creativecommons.org/licenses/by/4.0/. Reusers must credit tokenusage.dev, link the relevant note and licence, and identify changes. This package renames columns, converts dates and prices, adds documented nominal numeric and lab fields, and converts empty fields to missing values while preserving the source text.

Third-party source material keeps its own rights. Epoch AI's results retain their own CC BY 4.0 licence and credit to Epoch AI and the authors it names. The model-size help pages identify the relevant credits. Use `citation("tokenusagedata")` for the data records.

## Author(s)

Jack He, package author and maintainer.

Pistachio Technology Inc., data-record creator and compilation copyright holder.

## Source

https://tokenusage.dev/publications

doi:10.5281/zenodo.23552904

doi:10.5281/zenodo.23552975

doi:10.5281/zenodo.23553072

## Examples

```
data("token_prices", package = "tokenusagedata")
c(records = nrow(token_prices),
  model_lines = length(unique(token_prices$line)))
```

---

# context_windows *Published AI Context Windows*

---

## Description

Selected published context-window records from tokenusage.dev research note 003, retaining access restrictions, units and qualifications.

## Usage

```
context_windows
```

## Format

A data frame with 25 rows and 11 columns:

**date_text** Character. Verbatim date or observation field, including any date qualification.

**date** Date. The leading YYYY-MM-DD in `date_text`, parsed with `as.Date`; missing when the field does not start with a full date. The month-only MPT record is missing.

**model** Character. Verbatim model field, including the lab in parentheses.

**lab** Character. Text inside the last parenthesised group of `model`, without the parentheses. No lab names are normalized.

**window_text** Character. Verbatim published window label.

**window_tokens** Numeric. Nominal token count extracted from the first figure in `window_text`. Commas and a leading tilde are removed; suffixes k, M, B and T mean powers 3, 6, 9 and 12 of ten. A label with no figure would give a missing value.

**access** Character. Verbatim access description.

**availability** Character. Text before the first spaced middle dot in `access`: `Served`, `Limited` or `On paper`.

**qualification** Character. Verbatim explanation of the unit, access or other limitation.

**source_labels** Character. Published source labels in order, separated by `|` .

**source_urls** Character. Published source URLs in order, separated by spaces; archive targets and timestamps are retained.

## Details

The source is the October 2026 snapshot, researched through October 13. This is a selection of published records, not a continuous daily maximum. Training sequence lengths, input limits and input-plus-output totals remain different quantities. Tokenizers also differ, so equal token counts do not establish equal text capacity or effective recall. A served label includes public service or trained open weights; it does not assert unrestricted access to every hosted endpoint.

The numeric column is a reproducible extraction from the CSV label, not an independently measured effective limit. In particular, `65k` becomes 65000; Scout's `10M` becomes 10000000 rather than its configuration integer 10485760; and DeepSeek's `1M` becomes 1000000 rather than 1048576. A label describing a reduction from 2000000 to 1000000 gives 2000000, its first figure. Always read `window_text`, `access` and `qualification` alongside it.

Dates retain their published basis. GPT-3's derived date is its paper date, not its later API date. No day is invented for MPT. Grok 4.20's reduction is bounded by archive observations, and Grok 4.5's date retains its API-date qualification.

The CSV excludes the OpenRouter hosted listing for Llama 4 Scout and the LWM-Text-1M row containing a RULER result. The RULER comparison table is not included. The source selection also excludes research-only 10M and 100M claims, Pokee-Isaac's unreproduced claim, Mistral Large 4's unresolved window and early Kimi figures stated in Chinese characters. Excluded records are not reconstructed from the article or chart.

The source compilation is credited to tokenusage.dev and was deposited by Pistachio Technology Inc. It and this derived table are licensed under Creative Commons Attribution 4.0 International (CC BY 4.0): https://creativecommons.org/licenses/by/4.0/. Reusers must credit tokenusage.dev, link the note and licence, and identify changes. This package renames columns, adds the derived date, lab, nominal token count and availability, and converts empty fields to missing values. The original text is retained. Third-party source material keeps its own rights.

## Source

https://tokenusage.dev/publications/oct-26-context-windows

doi:10.5281/zenodo.23552975

Source file: `oct-26-context-windows.csv`, version 2026-10 snapshot.

## Examples

```
data("context_windows", package = "tokenusagedata")
table(context_windows$availability)
subset(context_windows, is.na(date),
       select = c(model, date_text, window_text))
```

---

`model_size_estimates` *Attributed AI Model Parameter Estimates*

---

## Description

Attributed third-party estimates and claims for closed models from tokenusage.dev research note 004. These are estimates, not lab disclosures.

## Usage

```
model_size_estimates
```

## Format

A data frame with nine rows and ten columns:

**date_text** Character. Verbatim model-date field, including grouped dates or other qualifications.

**date** Date. The leading YYYY-MM-DD in `date_text`, parsed with `as.Date`; missing if no full date leads the field. A field with two dates contributes only its leading date.

**model** Character. Verbatim model label, including grouped models.

**total_text** Character. Verbatim claimed or inferred total label, retaining approximation and evidence-type text.

**total_params** Numeric. Nominal total parameters extracted from the first figure in `total_text`. Commas and a leading tilde are removed; suffixes k, M, B and T mean powers 3, 6, 9 and 12 of ten. A label with no figure gives a missing value.

**active_text** Character. Verbatim active-per-token label, retaining any assumption.

**active_params** Numeric. Parameters active per token, extracted from `active_text` using the same first-figure rule. `Unknown` becomes a missing value.

**evidence** Character. Verbatim evidence and uncertainty, including confidence, prediction bands and alternative estimates.

**source_labels** Character. Published source labels in order, separated by `|` .

**source_urls** Character. Published source URLs in order, separated by spaces.

## Details

This is the closed-model selection in the October 2026 snapshot, researched through October 12. The source CSV omits the GPT-4 estimate derived from SemiAnalysis. It must not be restored from the note, another source or a chart.

The numeric columns are label extractions, not measurements or confidence bounds. The PaLM 2 active count assumes a dense model. Serving-economics estimates, knowledge-probe estimates, industry estimates and a CEO claim retain their different evidence types. The Grok conference transcription remains unverified. Its grouped row is not split into separately dated observations.

The source research also omits the withdrawn gpt-3.5-turbo figure, an unchecked o1 relay, unreleased Olympus and MAI-1 training targets, snippet-only GPT-5.6 Sol and GPT-6 Astra estimates, and models without usable totals. These estimates are not part of the note's disclosed model-size chart. They do not establish equivalent capability, measured inference speed or measured task cost.

The source compilation is credited to tokenusage.dev and was deposited by Pistachio Technology Inc. It and this derived table are licensed under Creative Commons Attribution 4.0 International (CC BY 4.0): https://creativecommons.org/licenses/by/4.0/. Reusers must credit tokenusage.dev, link the note and licence, and identify changes. This package renames columns, adds typed dates and nominal numeric counts, and converts empty fields to missing values. Original labels, uncertainty and source fields are retained.

Third-party source material keeps its own rights. Epoch AI's results retain their own CC BY 4.0 licence and credit to Epoch AI and the authors it names. In particular, credit Epoch AI and Ege Erdil for the GPT-4o and Claude 3.5 Sonnet serving-economics estimates, and retain their source URL. Removing a tilde for numeric extraction does not remove the estimate's uncertainty.

## Source

https://tokenusage.dev/publications/oct-26-model-sizes

doi:10.5281/zenodo.23553072

Source file: `oct-26-model-sizes-estimates.csv`, version 2026-10 snapshot.

## Examples

```
data("model_size_estimates", package = "tokenusagedata")
model_size_estimates[c("model", "total_text", "active_text")]
```

---

`model_sizes` *Published AI Model Parameter Counts*

---

## Description

Selected published total and active-per-token parameter counts from tokenusage.dev research note 004, with status and qualifications.

## Usage

```
model_sizes
```

## Format

A data frame with 28 rows and 11 columns:

**date_text** Character. Verbatim model-date field, including any qualification.

**date** Date. The leading YYYY-MM-DD in `date_text`, parsed with `as.Date`; missing if no full date leads the field.

**model** Character. Model portion before the spaced em dash in the source Model / lab field.

**lab** Character. Lab portion after the spaced em dash in the source Model / lab field. Joining `model` and `lab` with that separator reconstructs the source field exactly.

**total_text** Character. Verbatim published total-parameter label.

**total_params** Numeric. Total parameters extracted from the first figure in `total_text`. Commas and a leading tilde are removed; suffixes k, M, B and T mean powers 3, 6, 9 and 12 of ten. A label with no figure gives a missing value.

**active_text** Character. Verbatim active-per-token label.

**active_params** Numeric. Parameters active per token, extracted from `active_text` using the same first-figure rule. `Unknown` becomes a missing value.

**status** Character. Verbatim Status and qualifications field, including architecture, confidence and discrepancies.

**source_labels** Character. Published source labels in order, separated by `|` .

**source_urls** Character. Published source URLs in order, separated by spaces.

## Details

These 28 rows are the published-model selection in the October 2026 snapshot, researched through October 12. They are not the complete 118-model research appendix. No rows are omitted from this source CSV. Third-party closed-model estimates are separate in `model_size_estimates`.

Dates normally identify the earliest announcement naming a model. Later size disclosures, public weights and general availability remain separate qualifications. Weight status is fixed at the snapshot cutoff; an open-weight description can include a restrictive licence.

Totals retain lab headline figures. The numeric extraction does not resolve the differences recorded for DeepSeek-V3's MTP module, Kimi K2's later backbone count, Inkling's stored-weight count or Mistral Large 4's conflicting reports. Grok-1's active count remains qualified as a calculation from the published percentage. Unknown active counts remain missing. Read the numeric columns with their text labels and status.

The note's chart uses a disclosed dense-model record and five selected mixture-of-experts (MoE) examples with both total and active counts. Those MoE examples are not a size-record series. Switch-C is a qualified encoder-decoder research comparison outside the decoder-model sequence. Mistral Large 4 remains an API preview with weights pending. Parameter comparisons do not establish equivalent capability, measured inference speed or measured task cost.

The source compilation is credited to tokenusage.dev and was deposited by Pistachio Technology Inc. It and this derived table are licensed under Creative Commons Attribution 4.0 International (CC BY 4.0): https://creativecommons.org/licenses/by/4.0/. Reusers must credit tokenusage.dev, link the note and licence, and identify changes. This package renames columns,

splits the model/lab field losslessly, adds typed dates and nominal numeric counts, and converts empty fields to missing values. Original labels, qualifications and source fields are retained.

Third-party source material keeps its own rights. Epoch AI's results retain their own CC BY 4.0 licence and credit to Epoch AI and the authors it names. The Switch-C weight-status qualification cites Epoch AI; retain that credit and its source URL when reusing it.

## Source

https://tokenusage.dev/publications/oct-26-model-sizes

doi:10.5281/zenodo.23553072

Source file: `oct-26-model-sizes.csv`, version 2026-10 snapshot.

## Examples

```
data("model_sizes", package = "tokenusagedata")
known <- subset(model_sizes, !is.na(active_params))
head(data.frame(
  model = known$model,
  active_fraction = known$active_params / known$total_params
))
```

---

# token_prices *Historical API Token Prices*

---

## Description

Selected historical application programming interface (API) list prices from tokenusage.dev research note 002, in its October 2026 snapshot.

## Usage

```
token_prices
```

## Format

A data frame with 88 rows and 12 columns:

**line** Character. Selected model-line label, copied from the CSV.

**lab** Character. Model lab, copied from the CSV.

**date** Date. The leading YYYY-MM-DD in the source date field, parsed with `as.Date`; missing if no full date leads the field. All dates in this table have a full leading date.

**model** Character. Model represented by the line from this date, copied from the CSV.

**change** Character. Published event label: `start`, `step` or `relabel`.

**input_usd_per_mtok** Numeric. Published input price in US dollars per one million tokens, converted directly from its decimal CSV field.

**output_usd_per_mtok** Numeric. Published output price in US dollars per one million tokens, converted directly from its decimal CSV field.

**blended_usd_per_mtok** Numeric. Published, unrounded 3:1 blended price in US dollars per one million tokens, converted directly from its decimal CSV field.

**context** Character. Applicable context or prompt-length tier, copied verbatim; an empty field becomes `NA_character_`.

**caveats** Character. Published qualifications, copied verbatim; an empty field becomes `NA_character_`.

**note** Character. Published explanatory note, copied verbatim; an empty field becomes `NA_character_`.

**source_urls** Character. Source URLs in their published order, separated by spaces. Archive targets and timestamps are preserved.

## Details

The snapshot covers research through October 12, 2026. It contains nine selected model lines from six labs, not every model or the cheapest available API on each day. A successor replaces the represented model within a line even when the previous model remains available. The records are ordered by line and then date.

There are 44 starts, 15 price steps and 29 relabels. A relabel records a different model at the same rate. The opening davinci record on January 1, 2022 is an opening state for an already available model, not a launch event.

Prices are for uncached requests at the lowest prompt-length tier. Batch and cache discounts, free tiers, credits and negotiated rates are excluded. The selected DeepSeek peak prices exclude off-peak discounts. Labelled introductory and promotional prices count on their recorded dates. Some historical per-character prices were converted to tokens; their conversion assumptions remain in the caveats and notes.

The blended price is `(3 * input + output) / 4`. The published calculation uses integer micro-dollars before weighting. This table preserves the unrounded result, rather than the article's displayed cents. These are list-price comparisons, not measured task costs or comparisons at equal model quality.

The Meta line uses Together AI's hosted Llama 70B prices. Its July 2024 step also changes the default endpoint from Reference (FP16) to Turbo (FP8). The represented line ends on April 5, 2025, exclusively; the CSV contains no end-marker row. Exclude it from cross-line calculations on or after that date. This boundary is not a retirement date for Llama 3.3 API access. Grok enters on April 9, 2025.

Inferred dates, inferred prices, preview status, secondary sources and endpoint-specific prices remain qualified in the text. An archive capture date is not necessarily a verified effective date. GPT-4.1 is outside the selected OpenAI flagship sequence. Premium tiers are in `token_prices_premium`. No rows are omitted from this source CSV; narrative comparisons and article figures are not additional package data.

The source compilation is credited to tokenusage.dev and was deposited by Pistachio Technology Inc. It and this derived table are licensed under Creative Commons Attribution 4.0 International (CC BY 4.0): https://creativecommons.org/licenses/by/4.0/. Reusers must credit tokenusage.dev, link the note and licence, and identify changes. This package renames columns, converts dates and decimal prices to R types, and converts empty fields to missing values; it retains the published text and values. Third-party source material keeps its own rights.

## Source

https://tokenusage.dev/publications/oct-26-token-prices

doi:10.5281/zenodo.23552904

Source file: `oct-26-token-prices.csv`, version 2026-10 snapshot.

## Examples

```
data("token_prices", package = "tokenusagedata")
table(token_prices$change)
subset(token_prices, line == "OpenAI low-cost",
       select = c(date, model, blended_usd_per_mtok))
```

---

# token_prices_premium *Premium API Token Prices*

---

## Description

Four premium API tiers from tokenusage.dev research note 002, kept separate from its selected historical model lines.

## Usage

```
token_prices_premium
```

## Format

A data frame with four rows and seven columns:

**model** Character. Published premium-model or tier label.

**lab** Character. Published model-lab label.

**date** Date. The leading YYYY-MM-DD in the source date field, parsed with `as.Date`; missing if no full date leads the field. All four fields contain full dates.

**input_usd_per_mtok** Numeric. Published input price in US dollars per one million tokens, converted directly from the decimal CSV field.

**output_usd_per_mtok** Numeric. Published output price in US dollars per one million tokens, converted directly from the decimal CSV field.

**blended_usd_per_mtok** Numeric. Published, unrounded 3:1 blended price in US dollars per one million tokens, converted directly from the decimal CSV field.

**source_urls** Character. Published source URLs, separated by spaces if there is more than one.

## Details

These four tiers were priced above their lab's represented flagship line in the October 2026 snapshot. They do not form additional lines in `token_prices` and must not be added to calculations that use one price per represented line. The source snapshot covers research through October 12, 2026. These are historical records, not current price quotations.

Prices are for uncached requests at the lowest prompt-length tier. Batch and cache discounts, free tiers, credits and negotiated rates are excluded. The blended price is `(3 * input + output) / 4`; the source uses integer micro-dollars before weighting, and this table keeps its unrounded result. List prices do not measure task cost or establish equal model quality. The GPT-4.5 label retains its preview qualification. No rows are omitted from the premium source CSV.

The source compilation is credited to tokenusage.dev and was deposited by Pistachio Technology Inc. It and this derived table are licensed under Creative Commons Attribution 4.0 International (CC BY 4.0): https://creativecommons.org/licenses/by/4.0/. Reusers must credit tokenusage.dev, link the note and licence, and identify changes. This package renames columns and converts dates and decimal prices to R types. Empty fields would become missing values; there are no empty fields in this source file. Third-party source material keeps its own rights.

## Source

https://tokenusage.dev/publications/oct-26-token-prices

doi:10.5281/zenodo.23552904

Source file: `oct-26-token-prices-premium.csv`, version 2026-10 snapshot.

## Examples

```
data("token_prices_premium", package = "tokenusagedata")
token_prices_premium[
  order(token_prices_premium$blended_usd_per_mtok),
  c("model", "date", "blended_usd_per_mtok")
]
```

# Index

∗ **datasets**
context_windows, 3
model_size_estimates, 5
model_sizes, 6
token_prices, 8
token_prices_premium, 10
tokenusagedata-package, 2
∗ **package**
tokenusagedata-package, 2

context_windows, *2*, 3

model_size_estimates, *2*, 5, *7*
model_sizes, *2*, 6

token_prices, *2*, 8, *11*
token_prices_premium, *2*, *9*, 10
tokenusagedata
*(*tokenusagedata-package*)*, 2
tokenusagedata-package, 2