Overview of Time Series Foundation Models (TSFM)

Model	Open / Closed	Architecture	Pretraining dataset(s)	# Params	Output	Input
TimesFM (v1.0 / v2.0)	Open (Apache-2.0) (GitHub)	Decoder-only Transformer with patching (Google Research)	~100B real-world points incl. Google Trends & Wikipedia + synthetic (Google Research)	200M (v1.0), 500M (v2.0) (GitHub)	Point forecast; optional **experimental** 10-quantile heads (uncalibrated) (GitHub)	Univariate target; optional past/future covariates via external regressors (zero-shot) (GitHub)
Chronos	Open (Apache-2.0) (Hugging Face)	Encoder-decoder **T5** with value tokenization & de/quantization; autoregressive sampling (Hugging Face)	Large corpus of public TS + synthetic GP data (Hugging Face)	~8M / 20M / 46M / 200M / 710M (tiny→large) (Hugging Face)	Probabilistic (sampled trajectories) (Hugging Face)	Univariate target (single series per forecast) (Hugging Face)
Chronos-X	Open (code released)	Lightweight adapter modules: **IIB** (Input Injection Block) + **OIB** (Output	Inherits Chronos pretraining; adds covariate benchmark (32	Uses Chronos backbones (8M–710M); adapter overhead [small #	Probabilistic (inherits Chronos' sampled forecasts; OIB adjusts logits)	**Univariate target + covariates** (past & future)

		Injection Block) added to pretrained model; supports past & future covariates	synthetic datasets) and tests on 18 real-world datasets	extra params] (Hugging Face)		
Chronos-2	Open (Apache-2.0) (Hugging Face)	Encoder-only Transformer (T5-encoder-inspired) with patch embeddings + group attention for in-context learning across series/covariates	Subset of Chronos datasets + GIFT-Eval pretrain + large-scale synthetic multivariate/covariate data	~120M (base size)	Multi-step **quantile** forecasts (probabilistic)	Uni- or multivariate targets; supports **past-only** and **known-future** covariates (real & categorical)
Moirai	Open (Apache-2.0; code & weights) (GitHub)	Masked **encoder** Transformer; "any-variate attention" (handles variable # of variables) (arXiv)	**LOTSA**: 27B observations across 9 domains (arXiv)	14M / 91M / 311M (small/base/large)	Probabilistic (mixture distribution head) (arXiv)	Any-variate (uni/multi); supports covariates (arXiv)

Moirai-2	Open (cc-by-nc-4.0) (Hugging Face)	**Decoder-only** Transformer; multi-token generation; quantile loss; patch embeddings with missing-value flag; data filtering for non-forecastable series	GIFT-Eval (pretrain/train) + Chronos mixup (non-leaking subsets) + KernelSynth synthetic + internal Salesforce ops data	~12M params (estimated)	Probabilistic **quantile** forecasts (multi-step)	supports dynamic covariates via Uni2TS interface; more sizes likely forthcoming (HF collection)
Lag-Llama	Open (Apache-2.0) (Hugging Face)	Small decoder-only Transformer; uses **lags as covariates**	Diverse public TS corpora (paper)	**2.45M** (Hugging Face)	Probabilistic (distribution per step) (Hugging Face)	Univariate (lags; no exogenous by default)
Tiny Time Mixers (TTM)	Open (research weights; Granite variants Apache) (Hugging Face)	**TSMixer/MLP-Mixer** family with architectural tweaks (AP/DRS/RPT); multi-level channel mixing	Public datasets (Monash + LibCity), ~1B samples (pretrain)	**~1M+** (starts at ≈1M)	Point forecasts (paper focuses on accuracy/speed; no native probabilistic head reported)	**Multivariate**; can incorporate **exogenous** signals during fine-tuning (channel mixing)
TimeGPT-1	**Closed** (hosted	Transformer encoder-	>**100B** data points	NA	Probabilistic via	Uni- or multivariate;

	API) (Nixtla)	decoder specialized for TS	across many domains (public sources)		**conformal** prediction intervals; quantiles available via API (docs.nixtla.io, GitHub)	supports **exogenous variables** (past & future) via API (Nixtlaverse)

Notes:

- **TimesFM**: GitHub README (license, v1/v2 sizes, covariates & quantile heads) and Google Research blog (architecture, 100B pretraining mix). (GitHub, Google Research)
- **Chronos**: HF model card (license, sizes, data mix, sampling, univariate). (Hugging Face)
- **Chronos-X**: AISTATS 2025 paper (adapters, covariates, benchmarks) + Chronos card for backbone sizes. (Hugging Face)
- **Moirai**: ICML 2024 paper + Uni2TS GitHub (LOTSA size, any-variate, params; Apache-2.0). (arXiv, GitHub)
- **Lag-Llama**: HF model card (params, license [? double check], probabilistic output) + paper context. (Hugging Face)
- **Tiny Time Mixers**: arXiv (public data, ~1B samples, exogenous/channel mixing; starts at ~1M params); HF Granite/IBM cards for licensing context. (Hugging Face)
- **TimeGPT-1**: arXiv (architecture, >100B points; conformal intervals; param size unkown) + Nixtla docs (closed API, intervals, exogenous support). (Nixtla, docs.nixtla.io, Nixtlaverse)

Shortlist for macro + GDELT-style covariates + probabilistic output:

1. Moirai (Salesforce)

- **Strengths**: Handles uni/multivariate targets and dynamic covariates; features a **mixture-of-distributions** head for full predictive densities. (arXiv)
- **Scale & data**: Pretrained on **LOTSA** (~27B observations, 9 domains); available in **Small/Base/Large (14M/91M/311M)** sizes for flexibility. (Salesforce, Hugging Face)
- **Notes**: Weights are released under **CC-BY-NC-4.0** (non-commercial); code in Uni2TS. (Hugging Face, GitHub)
- **Notes**: uni2ts can be added to environment to ease use of model

2. Chronos (Amazon)

- T5-family **encoder–decoder** LM; values are **tokenized**; forecasts are **probabilistic** via sampling multiple futures. (Hugging Face, GitHub)
- **Caveat**: **No native covariates**; designed for **univariate targets** only. Participants can add covariate effects through external models, residual forecasting, etc. (Hugging Face)

3. Chronos-2 (Amazon)

- Natively ingests covariates and multivariate signals.
- Outputs quantiles directly (no sampling loop).
- Much simpler to deploy in a classical forecasting pipeline (X, y, horizons).

4. Moirai-2 (Salesforce)

- Decoder-only TS FM with direct multi-step **quantile** outputs, patch embeddings (with missing-value flags), and stronger data hygiene (filters non-forecastable series).
- Trains on GIFT-Eval + curated Chronos-style mixes + KernelSynth synthetic; released under **CC-BY-NC-4.0** with a small checkpoint prioritized for efficiency
- Good fit for lightweight deployment with native uncertainty