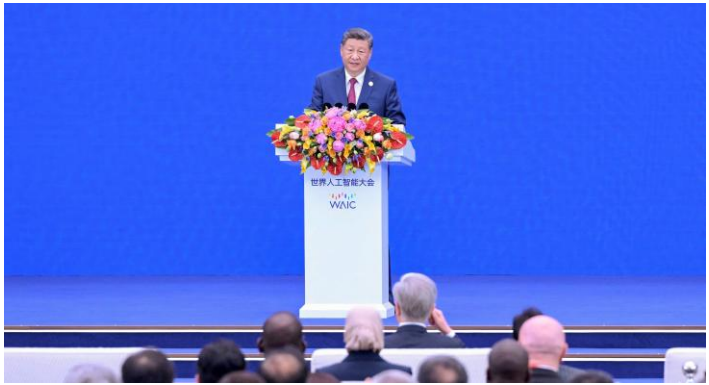
通義千問(Qwen)의 全球 人工智能 覇權 戰略 및 技術 生態系 分析

부산대학교 BAELAB
석사과정 박택현
2026年 07月 27日

0. 世界人工智能大會(WAIC) 習近平 국가주석 基調演說

- 2026년 7월 17일, 상하이 세계인공지능대회(WAIC)에서 習近平 국가주석은 중국 AI 전략의 방향을 네 가지 원칙을 제시함.
  - 개방·혁신: 오픈소스, 협업, 성과 공유를 통한 AI 확산
  - 안전·통제가능성: 기술 발전과 위험 관리의 병행
  - 포용·공유: 개발도상국의 AI 역량 강화 및 접근성 확대
  - 화충공제(和衷共濟): 다자 협력을 통한 공정하고 합리적인 글로벌 AI 거버넌스 구축
- 이를 통해 중국은 자국 LLM 생태계를 단순한 기술 자립의 수단을 넘어, 제3국이 선택할 수 있는 미국 중심의 AI질서에 대응하는 독자적 AI 기술권으로 확장하고자 함.



1. 中國의 Foundation LLM 生態系

- 중국 AI 생태계는 반도체(華爲(Huáwéi) Asend 950C)·메모리(長鑫存儲(CXMT) HBM) 등 하드웨어, 컴퓨터 아키텍처(RISC-V)·AI 프레임워크(PaddlePaddle) 등 소프트웨어, LLM·비전·오디오를 포괄하는 응용 모델까지 자체 공급망을 빠르게 구축 중임.
- LLM 분야 역시 대규모 사전학습부터 추론·배포까지 독자적으로 운영할 역량을 확보하고 있음. 중국 모델 기업들은 모델을 오픈 소스로 공개하며 성능 경쟁을 넘어 개발 도구·데이터 형식·응용 방식 전반의 기술 표준을 넓혀가는 중임.
- 그 결과 Qwen, GLM, Kimi 등 중국계 오픈 모델은 국내외 연구 및 서비스 개발의 기반 모델로 빠르게 채택되고 있음. 중국 LLM 생태계는 미국 모델의 단순한 대체재가 아니라, 제3국이 선택할 수 있는 독립적 AI 기술권으로 자리 잡는 과정에 있음.


1. 中國의 Foundation LLM 生態系

- 중국 LLM 생태계에는 Moonshot AI의 Kimi, DeepSeek, Zhipu AI(Z.ai)의 GLM, MiniMax의 MiniMax, Tencent의 Hunyuan, Xiaomi의 MiMo, Meituan의 LongCat, Baidu의 ERNIE 등 다양한 기업과 모델이 경쟁 중임.
- Qwen은 소형 모델부터 대규모 모델까지 폭넓은 모델군을 공개해, 연구·경량화·상용 배포 등 다양한 환경에서 활용되고 있음. 특히 모델 가중치뿐 아니라 기술 보고서, 학습·추론 도구, 에이전트 프레임워크 등을 폭넓게 공개하며 오픈소스 생태계 확산에 기여해왔음.
- 기존의 Kimi, Z.ai 같은 경우, Agent AI에만 집중에서 아키텍처와 개발을 했다면, Qwen 같은 경우 Vision 인식, Embedding, Video Generation 등 Agent 뿐만 아니라 Pre-training 한 LLM 을 활용한 다양한 Multimodal Task를 학습하는 방법을 을 공개해왔음.

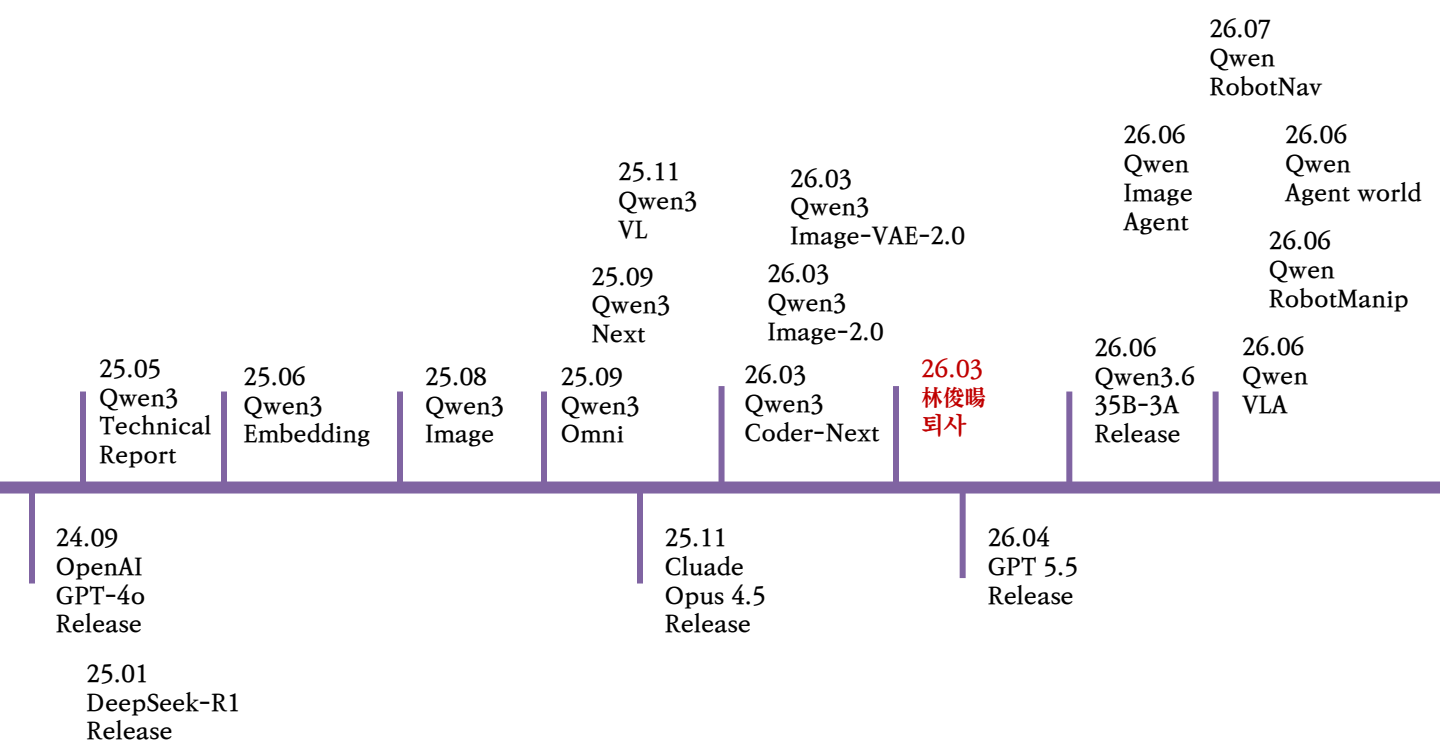
2. 通義千問 (Tongyi Qianwen)

- 린쥔양(林俊暘, 1993生)은 北京大學에서 영어학 학사, 계산언어학 석사 학위를 마친 뒤,, 2019년 Alibaba DAMO Academy에 선임 알고리즘 엔지니어로 합류함. 이후 텍스트 중심의 AliceMind 계열과 멀티모달 사전학습을 지향한 M6 계열 연구가 병행되던 시기에, M6·OFA·Chinese CLIP 등 언어와 비전을 결합하는 대규모 멀티모달 모델 연구를 주도함.
- GPT-3 공개 이후 Alibaba 역시 대규모 사전학습 모델을 핵심 전략으로 전환했고, M6는 2020년부터 100B 파라미터 규모로 확장되며 중국어 멀티모달 사전학습 모델의 기반을 마련함. 이후 2022년 말 DAMO Academy의 언어·비전 연구팀이 Alibaba Cloud로 통합되어 通義랩 설립됐고, 林俊暘 은 Qwen의 기술 책임자로서 모델 개발을 이끌게 됨.
- 이후 2023년 4월, Qwen LLM을 발표하면서 0.6B ~ 236B 부터 다양한 모델들을 Open Source로 공개하면서, SLM의 실험 논문들의 표준이 됨.

2. 通義千問 (Tongyi Qianwen)

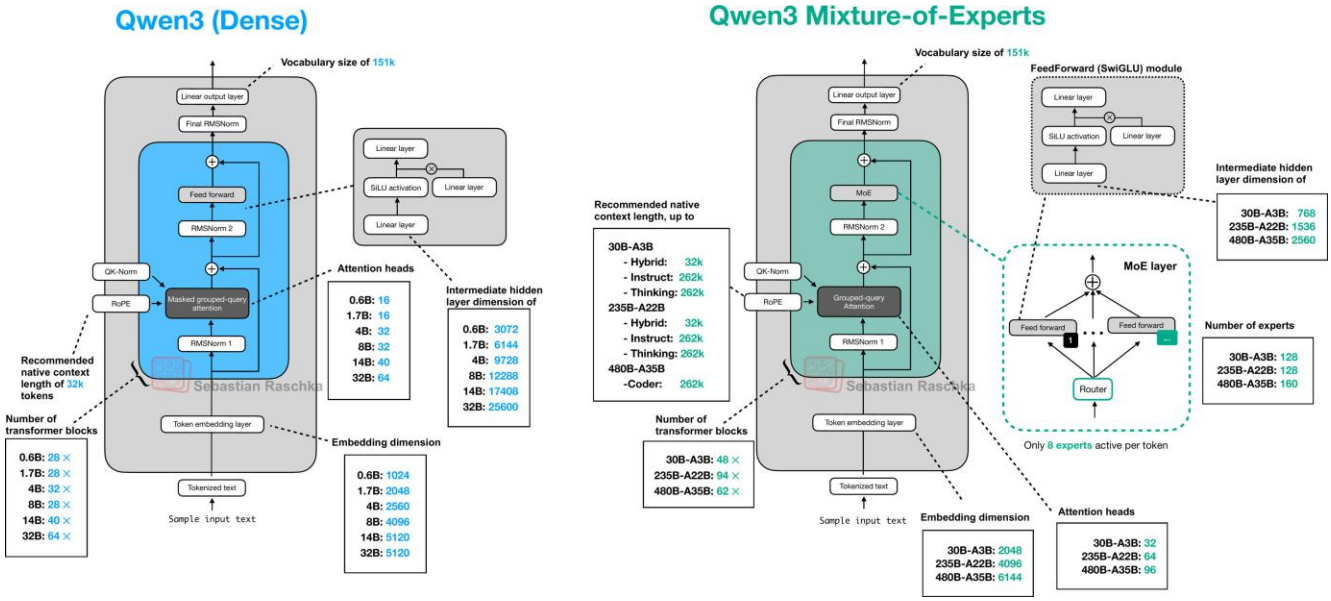
- 하지만 2026년 3월 4일 퇴사. 그리고 그가 이후에 X 에 마지막으로 기고한 글. "*From "Reasoning" Thinking to "Agentic" Thinking*"
- O1 과 DeepSeek-R1 이후의 시대는 Reasoning의 시대였음. GRPO와 GSPO를 기반으로 Reasoning을 수행하여 수학문제, 코드를 검토하는 형태로 나아가고 있었음.
- 하지만, 긴 추론이 곧 높은 지능으로 나아갈 수 없음. 일정 수준이상의 생각 깊이에 다다르면 더 이상 smart 하지 않음. (실제로 Opus 5의 경우 Medium이 xHigh 보다 성능이 높음)
- 핵심은 Agentic Thinking. 행동을 하고, 행동의 결과를 관측하고 이를 Closed Loop로 수행하는 것. 이를 통해 생각의 결과와 행동까지 언제 해야 할지 결정하는 것 까지 사용해야 함.
- 향후 성능은 모델 가중치만으로 결정되지 않음. MCP,브라우저,터미널,시뮬레이터,메모리,평가 환경을 결합하는 Harness Engineering과, 실제 환경 피드백을 활용하는 Agentic RL이 핵심 경쟁력이 될 것임.

2. 通義千問 (Tongyi Qianwen)



3. Qwen3 Technical Report (25.05)

- Dense, MoE 모델 전부 전작(Qwen2.5)와 동일하게 GQA, SwiGLU, RoPE, AdamW 를 사용함. 다만 QK-Norm& QKV bias 제거를 Qwen3 부터 도입. MoE의 경우, 전작 Qwen2.5-MoE 와 다르게 Shared Experts 를 배제하고 128개중 단순 8개의 Activated Expert를 사용하게 함. + Global-bath load balancing loss로 Expert 사용량을 균형 있게 유지 (다만, 최근 Kimi 나 GLM, DeepSeek의 MoE 경향성과 이후 Next에서 다시 Shared를 적용한 것을 보면 Shared Expert를 쓰는 게 좋은 선택이라고 볼 수 있을 듯)



3. Qwen3 Technical Report (25.05)

- Pre-training Stage

General Stage

4096 tokens 를 기준으로
30T 토큰을 학습.
(119 Languages & dialects)

Reasoning Stage

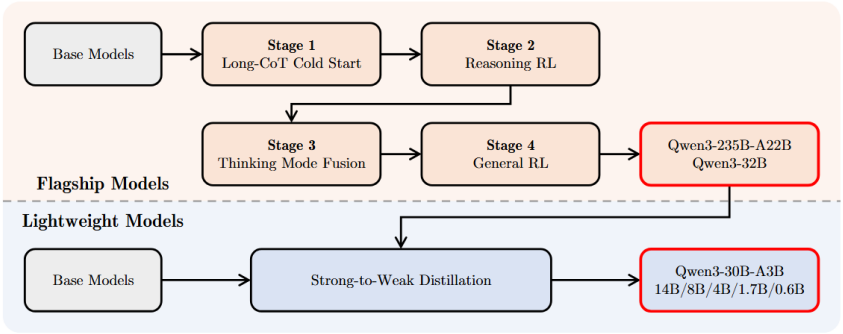
STEM, coding, reasoning, and synthetic data 의 고품질 데이터를 이후에 학습.
(4096tokens 를 기준으로 5T token 학습)

Long Context Stage

1. ABF technique 으로 RoPE 주파수 범위를 확장 (4k -> 32k 확장) 및 학습.
2. YaRN 기반으로 Inference 단계에서 Extrapolation 수행 (32k -> 128k)
3. Dual Chunk Attention 을 통해 Inference 단계에서 Attention 효율화 (32k -> 128k)

3. Qwen3 Technical Report (25.05)

- Post-Training
  - Qwen3는 기존 LLM과 달리 Thinking Mode와 Non-Thinking Mode를 하나의 모델에 통합해 학습하는 방식을 제시함. Chat Template에서 /think와 /no_think 지시어를 사용해, 상황에 따라 깊은 추론과 빠른 응답을 선택할 수 있도록 함.(최근 LLM 모델의 생각의 양을 조절하는 것도 이 방법의 파생)
  - 학습 과정은 일반적인 RL 학습과 유사하게 구성됨. Stage 1에서는 장문 CoT 데이터를 활용한 SFT로 기본적인 추론 능력을 학습하고, Stage 2에서는 수학·코딩 과업을 중심으로 Reasoning RL을 수행함. Stage 3에서는 Thinking Mode와 Non-Thinking Mode의 데이터를 함께 활용하는 SFT를 통해, 하나의 모델이 두 가지 응답 방식을 모두 수행하도록 학습함.
  - 이후 Stage 4에서는 수학·코딩 외의 일반 과업에서도 사용할 수 있도록 General RL을 수행함. 또한 소형 모델은 대형 모델의 출력을 활용한 Distillation으로 학습하며, Off-policy와 On-policy 방식을 함께 적용해 추론 능력을 이전함.

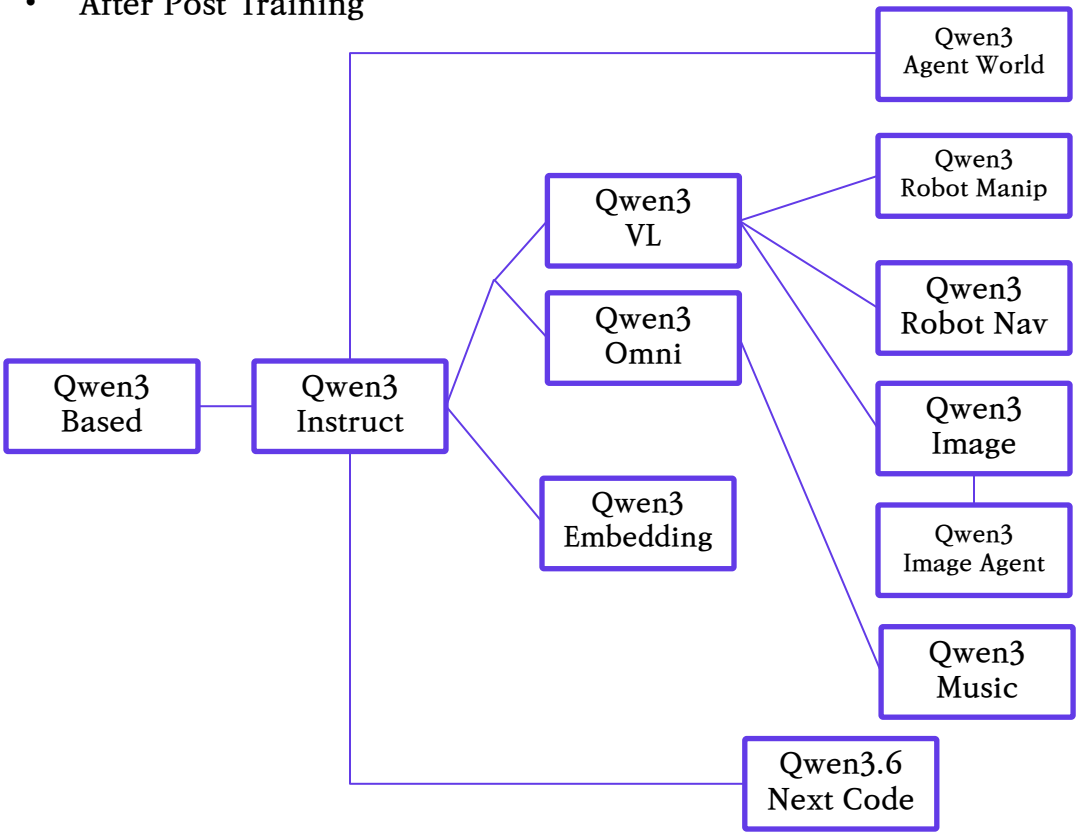


Thinking Mode	Non-Thinking Mode
<|im_start|>user {query} /think<|im_end|> <|im_start|>assistant <think> {thinking_content} </think> {response}<|im_end|>	<|im_start|>user {query} /no_think<|im_end|> <|im_start|>assistant <think> </think> {response}<|im_end|>

Figure 1: Post-training pipeline of the Qwen3 series models.

3. Qwen3 Technical Report (25.05)

- After Post Training



4. Qwen3 Embedding

- Qwen3 Embedding: Advancing Text Embedding and Reranking Through Foundation Models
  - Qwen3의 소형모델(0.6,4,8B)들을 기반으로 Embedding & Reranker 모델을 생성하는 방법론.
  - Embedding 모델은 Query를 `Instruct: {Task}\n Query: {Query}` 형식으로 입력하고, Document는 별도 instruction 없이 입력함. 각 입력의 마지막 [EOS] 토큰 hidden state를 추출한 뒤 L2 정규화하여 임베딩 벡터로 사용하며, Query-Document 간 cosine similarity를 통해 검색을 수행함. 학습은 대규모 약지도 Contrastive Pre-training, 고품질 라벨 데이터 기반 Supervised Fine-tuning, 그리고 여러 후보 모델을 결합하는 Model Merging의 3단계로 이루어짐.
  - Reranker는 Query와 Document를 하나의 입력으로 결합하는 Cross Encoder 방식임. `<Instruct>:<Query>:<Document>`를 함께 입력한 뒤, 문서가 질의 요구사항을 충족하는지에 대한 yes와 no 토큰의 확률을 비교하여 relevance score를 산출함.

$$\text{score}(q,d) = \frac{e^{P(\text{yes}|I,q,d)}}{e^{P(\text{yes}|I,q,d)} + e^{P(\text{no}|I,q,d)}}$$

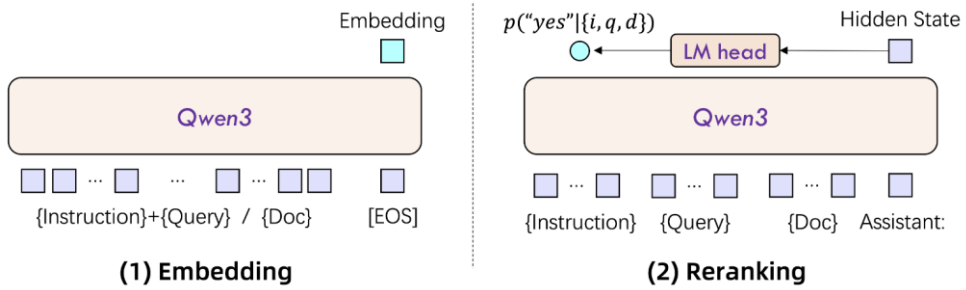


Figure 1: Model architecture of Qwen3-Embedding (left) and Qwen3-Reranker (right).

4. Qwen3 Embedding

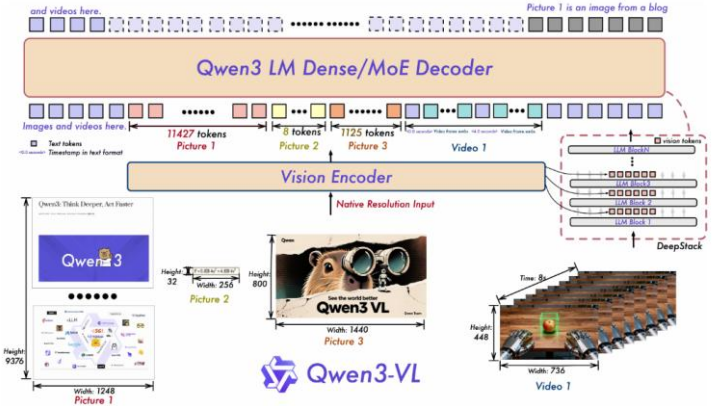
- Qwen3 Embedding: Advancing Text Embedding and Reranking Through Foundation Models

```
{Instruction} {Query}<|endoftext|>
```

```
<|im_start|>system
Judge whether the Document meets the requirements based on the Query and the
↪  Instruct provided. Note that the answer can only be "yes" or
↪  "no".<|im_end|>
<|im_start|>user
<Instruct>: {Instruction}
<Query>: {Query}
<Document>: {Document}<|im_end|>
<|im_start|>assistant
<think>\n\n</think>\n\n
```

5. Qwen3-VL (2025.12)

- Qwen3-VL Technical Report
  - Qwen3-VL은 Vision Encoder,MLP Merger,Qwen3 LLM으로 구성된 VLM임. Vision Encoder로는 동적 해상도를 지원하도록 SigLIP-2를 사용하며, 이미지·영상 입력을 가변 길이의 visual token으로 변환해 Qwen3 LLM에 전달함.
  - 전작 대비 핵심 변화는 세 가지임. 첫째, 시간·가로·세로 위치 정보를 주파수 대역 전체에 고르게 배치하는 Interleaved-MRoPE를 도입해 장문 영상의 공간·시간 정보를 더 안정적으로 표현함. 둘째, Vision Encoder의 여러 layer에서 나온 feature를 초기 LLM layer에 주입하는 DeepStack을 적용해 저수준 시각 정보와 고수준 의미 정보를 함께 활용함.
  - 셋째, 기존의 시간 위치 인코딩(T-RoPE) 대신 영상 frame group 앞에 〈3.0 seconds〉와 같은 텍스트 timestamp token을 넣음. 이를 통해 모델이 영상 속 사건의 시간적 위치를 직접 이해하게 하며, Video Grounding·Dense Captioning과 같은 시간 단위 영 상 이해 성능을 높이고자 함.



5. Qwen3-VL (2025.12)

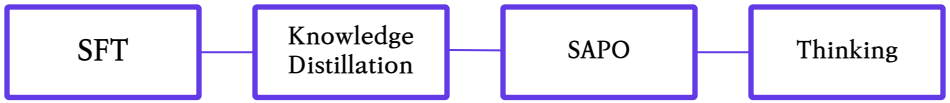
- Pre-Training
  - 학습 데이터는 이미지 캡션이 존재하는 일반 데이터 & OCR 기반 데이터 & 멀티모달 추론 데이터 & GUI 기반의 Agent Data 이며
  - 이 데이터들을 기반으로 총 4단계를 통해 학습을 수행함.
    - S0: Vision Encoder와 LLM 사이의 Align을 진행 (인코더, LLM은 얼리고 사이 MLP Merger만 학습)
    - S1: 이후 전체 파라미터에 대해 1T 규모로 학습.
    - S2: 8k->32k 로 시퀀스를 확장하고 영상 + 에이전트 데이터 학습
    - S3: 시퀀스를 256k로 확장해서 장문영상 장문 문서 기준으로 직접 학습

Table 1: Training setup and hyperparameters across different stages for Qwen3-VL.

Stage	Objective	Training	Token Budget	Sequence Length
S0	Vision-Language Alignment	Merger	67B	8,192
S1	Multimodal Pre-Training	All	~1T	8,192
S2	Long-Context Pre-Training	All	~1T	32,768
S3	Ultra-Long-Context Adaptation	All	100B	262,144

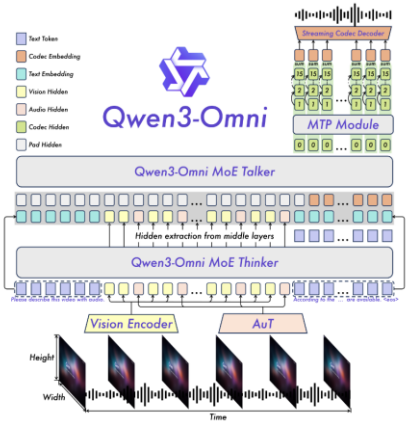
5. Qwen3-VL (2025.12)

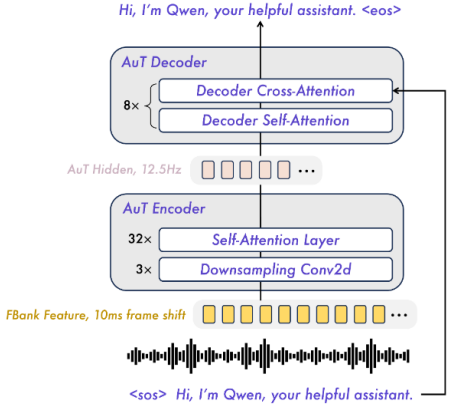
- Post-Training
  - SFT: Pre-training 이후, Text-only, Image-Text, Video-Text, 데이터를 결합해 SFT를 수행함. 일반 Instruct 모델용 데이터와 Long-CoT 기반 Thinking 모델용 데이터를 분리해 학습하며, 32K context 학습 이후 장문 문서·장문 영상 데이터를 중심으로 256K context까지 확장함.
  - Knowledge Distillation: 이후 강력한 Teacher 모델의 능력을 소형 Student 모델로 이전하는 Strong-to-Weak Distillation을 수행함. 이 단계는 Text-only 데이터로 LLM backbone을 Fine-tuning하지만, 텍스트 추론 능력의 향상이 멀티모달 추론 과업에도 유의미하게 전이됨을 보임. Off-policy Distillation과 On-policy Distillation을 함께 적용함.
  - SAPO: 추론 RL에서는 수학·코딩·논리·Visual Grounding·Visual Puzzle 등 정답 검증이 가능한 텍스트·멀티모달 과업을 중심으로 SAPO 를 적용함. 이후 General RL에서는 VQA, Image Captioning, OCR, Instruction Following에 대해 Rule-based Reward와 Model-based Reward를 함께 사용해 일반화 성능과 응답 품질을 개선함.
  - Thinking: 단순히 이미지 입력만으로 추론하는 방식이 아니라, 시각 에이전트가 Think → Act → 환경 피드백 분석 → Answer의 과정을 수행하도록 학습함.



6. Qwen3-Omni (2025.09)

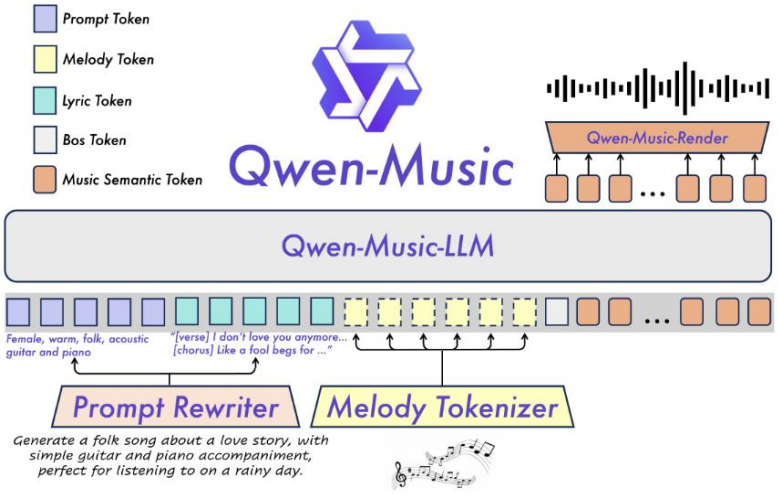
- Qwen3-Omni Technical Report
  - Qwen3-Omni는 Qwen3 LLM에 Qwen3-VL의 Vision Encoder와 자체 사전학습한 AuT Audio Encoder를 결합하고, Encoder Alignment →전체 모달리티 학습(약 2T tokens) → 32K Long-context 학습의 순서로 확장한 Omni-modal 모델임.
  - 출력은 멀티모달 이해·텍스트 생성을 담당하는 Thinker(30B-A3B)와, 음성 생성을 담당하는 Talker(3B-A0.3B)로 분리해 기존 Qwen3의 추론 능력과 실시간 음성 응답을 결합함.
  - Talker는 Thinker가 생성한 텍스트와 멀티모달 표현을 받아, 음성 waveform이 아닌 discrete speech codec token을 생성하는 MoE Transformer임. 매 시점 첫 번째 codec token은 자기회귀적으로 생성하고, MTP(Multi-Token Prediction) module이 이를 바탕으로 나머지 codebook token을 병렬 예측함. 이후 Code2Wav Causal ConvNet이 생성된 codec token을 즉시 waveform으로 변환하여, 전체 문장이 끝나기 전에도 실시간으로 음성을 출력할 수 있음.





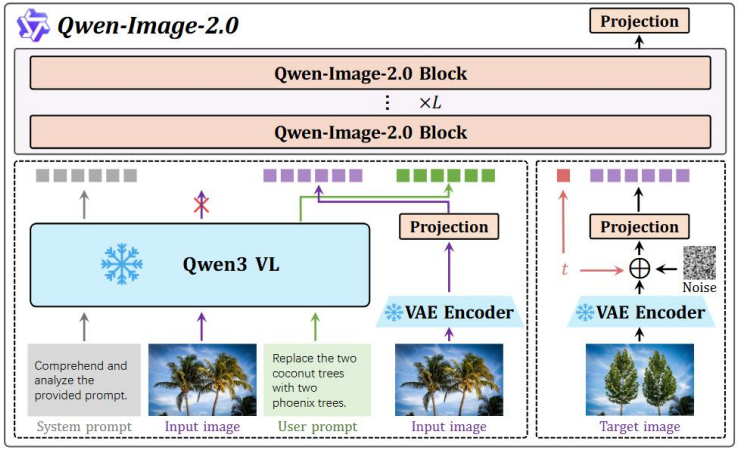
7. Qwen3-Music (2026.07)

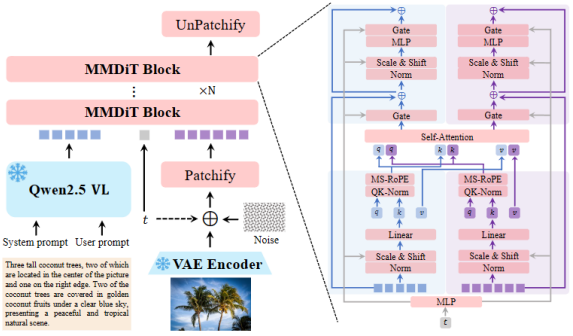
- Qwen-Music Technical Report
  - Qwen-Music은 Qwen3.5-Omni의 3B Dense LLM을 backbone으로 초기화해, 음악 생성으로 확장한 모델임.
  - 음악은 Conformer 기반 Music-Tokenizer가 25Hz Music Semantic Token으로 변환함. 생성된 Music Semantic Token은 Qwen-Music-Render가 고음질 stereo waveform으로 변환함. Render는 Spec-VAE와 DiT(Diffusion Transformer), Band-Mode Refiner를 사용.
  - 학습은 500만 시간 이상 다국어 음악 데이터에 대한 quality-aware Pre-training 이후, Supervised Initialization → Offline DPO → Online GSPO 순으로 진행함.



8. Qwen3-Image (2026.03)

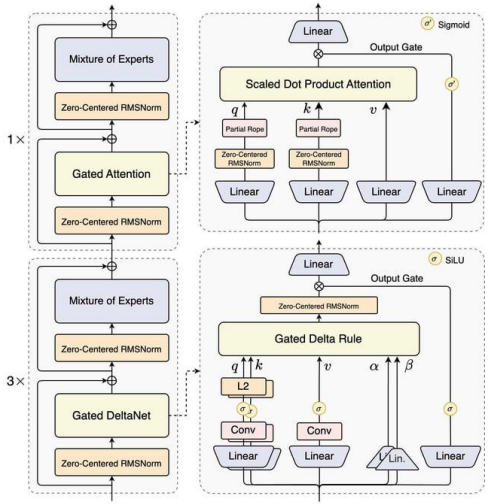
- Qwen-Image Technical Report
  - Qwen2.5-VL-7B를 동결된 Text Encoder로 사용하고, 20B MMDiT(Multimodal Diffusion Transformer)를 이미지 생성 backbone으로 사용하는 모델임.
  - Text Encoder의 표현과 VAE의 image latent를 MMDiT에서 함께 처리하며, 텍스트와 이미지의 위치 정보를 정렬하기 위해 MSRoPE(Multimodal Scalable RoPE)를 사용함. 이를 통해 고해상도 이미지 생성과 복잡한 텍스트 렌더링, 이미지 편집 성능을 강화함.
  - 학습은 Flow Matching 기반 Pre-training으로 시작하고, 이후 SFT → DPO → GRPO 순으로 진행함. SFT는 지시사항과 이미지 편집 Task를 정렬하고, DPO는 선호 데이터로 이미지 품질을 개선하며, GRPO는 Reward 기반의 Online RL을 통해 텍스트 정확성·미학·프롬프트 준수 성능을 높임.

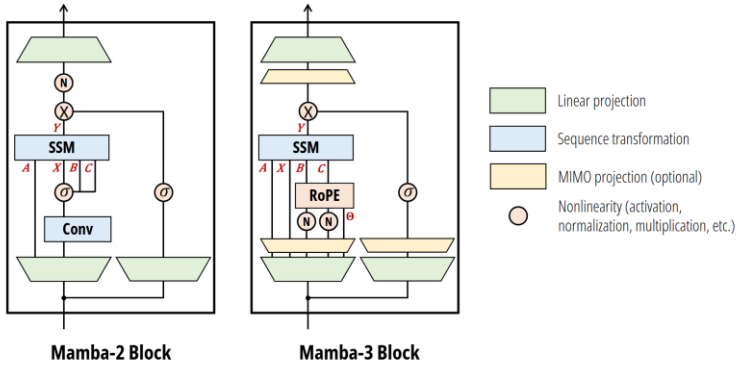




9. Qwen3-Next (2025.10)

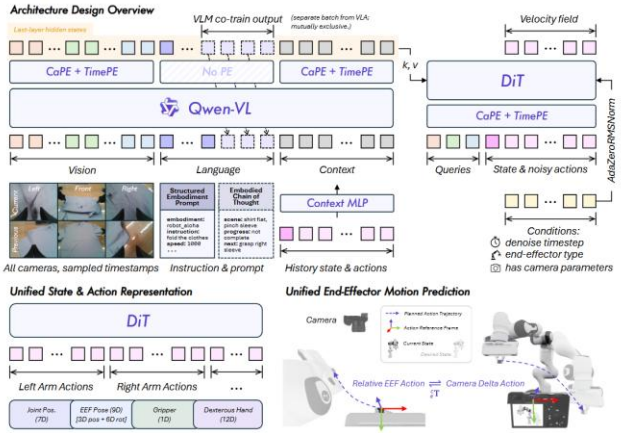
- Qwen3-Next
  - Agent 시스템 + CoT(GSPO,GRPO) 가 확산되면서 LLM이 처리해야 하는 Context 사이즈가 늘고 있음. 이 환경에서 Attention 비용을 줄이기 위해 Linear Attention 계열들(Kimi Delta Network, Gated Delta Network, Mamba) 가 LLM에 섞이는 Hybrid Attention 구조가 사용되고 있음. (Solar Open2, KIMI-K3, Nemotron,Ling)
  - Qwen3-Next는 Gated DeltaNet과 Gated Attention을 결합하고, Sparse MoE를 적용해 장문 Context 환경에서 성능과 추론 효율을 함께 확보하고 있음. Qwen 3.5 부터는 Next 구조를 도입해서 사용.





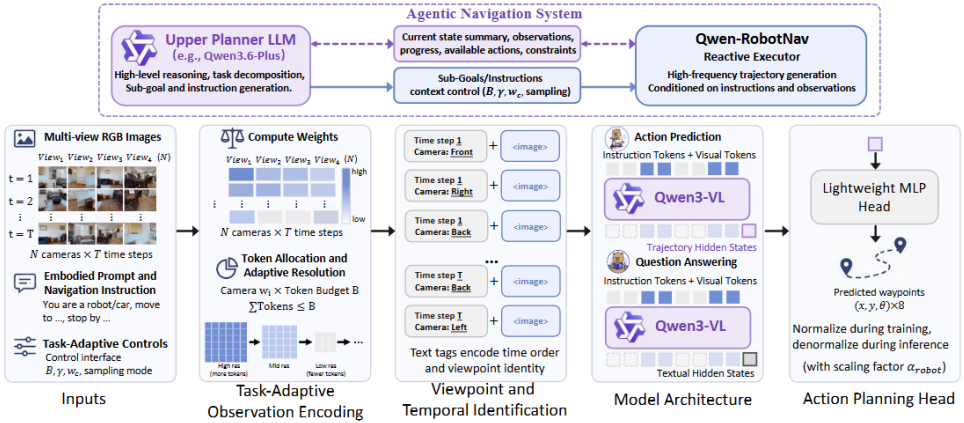
10. Qwen3-RobotManip (2026.07)

- Qwen-RobotManip Technical Report: Alignment Unlocks Scale for Robotic Manipulation Foundation Models
  - Qwen-RobotManip은 Qwen3.5-4B VL backbone에 Flow Matching 기반 DiT Action Head를 결합한 VLA 모델임. 다중 카메라 영상, 자연어 지시, 로봇 상태·과거 행동을 Qwen-VL이 함께 인코딩하고, DiT가 연속적인 로봇 행동 trajectory를 생성함.
  - 서로 다른 로봇의 데이터를 함께 학습하기 위해, 관절·End-effector pose·Gripper·손 동작을 공통의 80차원 State-Action vector로 정렬함. 또한 End-effector action을 로봇 base 좌표계가 아닌 Camera-frame delta pose로 표현해, 화면에서 유사하게 보이는 움직임이 action space에서도 유사하도록 만듦.
  - CaPE(Camera Positional Encoding)는 카메라의 내·외부 파라미터를 DiT의 Cross-Attention에 주입하는 방식임. 따라서 모델이 여러 카메라 view와 로봇 자세의 기하학적 관계를 고려해 action을 생성할 수 있음



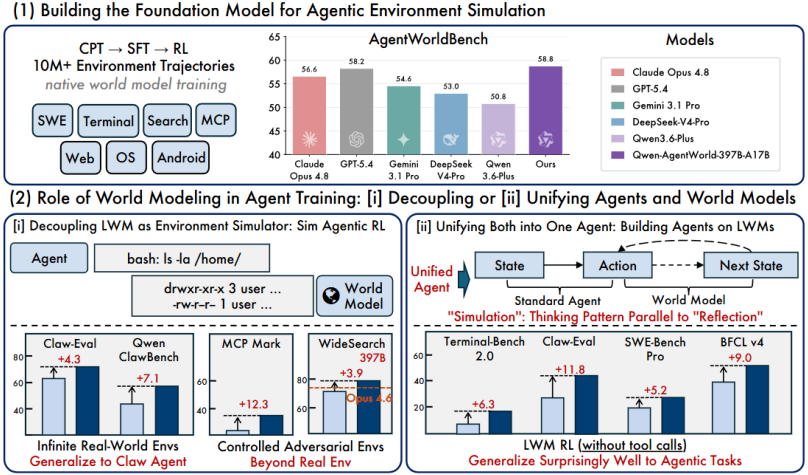
11. Qwen3-RobotNav (2026.07)

- Qwen-RobotNav Technical Report: A Scalable Navigation Model Designed for an Agentic Navigation System
  - Manipulate 를 넘어서 Robot Navigation 을 수행함.
  - 기존 NavFoM, Abot-N0 와 다르게 전용 Robot Navigation Architecture를 설계하지 않고, Qwen3-VL 을 Backbone으로 사용하여 Context Modeling으로 수행하게 만듦
  - 상위 Planner로 대형모델을 기반으로 Subgoal을 생성하고, 이후 Qwen-RobotNav로 사용하게 만듦
  - 학습은 Lightweight MLP Head만을 두고, 이걸 기반으로 웨이포인트를 생성하게 만듦



12. Qwen-AgentWorld (2026.07)

- Qwen-AgentWorld: Language World Models for General Agents
  - Language World Model(AI Agent가 현재행동과 상태를 기반으로 다음 상태를 예측하는 언어 모델)
  - MCP, SWE, Search, Terminal, Web, Android, OS 와 같은 언어 기반의 World 시스템을 설정하고, 미래를 예측할 수 있도록 함.
  - 학습은 Qwen 3.5 backbone에 CPT(Continual Pretraining), SFT, RL(GSPO)을 통해서 진행하고 있음.



13. Future of Qwen

- Qwen은 Vision, Audio, Image, Robot 등 다양한 Modality를 별도 모델로 분리하기보다, LLM을 공통 backbone으로 활용해 통합하는 방향을 보여주고 있음. Image Generation, Music Generation, Robot Manipulation에서도 기존 Qwen LLM Representation Space를 재사용하며, Agentic Foundation Model을 각 응용 도메인으로 확장하는 전략을 취하고 있음.
- Agent 시스템을 위한 Long Context 처리를 위해 Qwen3 에 Linear Attention , MTP 를 결합하는 형태로 계속해서 시스템을 최적화 시키고 있음.
- 최근 논문들에서 공통적으로 보이는 방향은 Agent가 행동하기 전 미래 상태와 행동 결과를 예측할 수 있도록 World Model을 구축하는 것임. 즉 Qwen의 다음 단계는 언어를 생성하는 모델을 넘어, 언어·시각·행동 데이터를 통합해 환경을 이해하고 시뮬레이션하며 행동하는 World Model로 확장되는 과정이라고 볼 수 있음.

14. Future of Port Logistic

- Qwen에서 확인한 AI의 발전 방향은 복잡한 문제를 추론하는 CoT, 외부 시스템을 활용해 계획·실행하는 Agent, 그리고 행동 이후의 환경 변화를 예측하는 World Model로 확장되는 흐름임.
- 해운·항만 분야에서도 AI가 나아가야 할 방향은 현재 상태를 보여주는 관제나 단일 과업 최적화를 넘어, 운영 결정 이후의 미래를 시뮬레이션하고 예측하는 World Model이라고 볼 수 있음. 항만은 선박 지연, 기상 악화, 장비 고장, 화물 반입 변동, 글로벌 공급망 충격 등 비정형적이고 불확실한 변수가 서로 연쇄적으로 영향을 미치는 환경임.
- 따라서 핵심 과제는 “현재 항만이 어떤 상태인가”보다, “지금 선석·장비·야드 운영을 바꾸면 이후 혼잡·대기시간·탄소배출·안전 리스크가 어떻게 변하는가”를 어떻게 예측할 것인가에 있음. World Model은 선박·선석·야드·장비·육상운송·기상 데이터를 통합한 가상 항만 환경에서 다양한 운영 시나리오를 사전에 검증하는 기반이 될 수 있음.
- 이러한 방향성은 단순히 정보를 검색하고 도구를 호출하는 LLM Agentic System을 넘어, 언어·시각·시계열·운영 데이터를 바탕으로 환경의 변화를 이해하고 미래 상태를 생성하는 Language World Model로의 확장이라고 볼 수 있음.

Q&A