Open Source Licensing ... for Success

Malcolm Bain

(Across Legal - Barcelona)

CURIOSS and Ireland Open Innovation Summit

February 2026

"10 simple rules for making your software outlive your job" (CURIOSS et al., 2025)

- Rule 1: stay within the law
    - See if you can release your code
    - Respect other open source licenses
    - Avoid incompatibilities
- Rule 4: choose an open license (if you can)
    - Put a (open source) license on your code
    - Creative commons licenses…. Are for documentation and data!

Open source and Licensing

- Open source is a collaborative software development model
- The license is only an instrument / tool to enable the model
- Successful licensing is not the objective: successful collaborative development is.
- So license choice is about enabling better and more collaborative development… and take up.
- This may mean to make your license fit your technology stack and your community

Importance of getting licensing right

- All code needs and license (and data, and AI models)
- Establishes collaboration mechanism with community
- Critical factor in commercial open source ventures
  - See Linux foundation: "State of Global Open Source" and "State of Commerical Open Source" research
- Key technical, legal and community considerations

Licenses 101

- Many open source licenses….
- 6-8 main licenses
  - "public licenses" GPL, LGPL, MPL, EPL
  - "permissive": Apache, MIT, BSD, EDL
- General categorisation between permissive and copyleft
  - Historically, strong leaning towards copyleft licenses (GPL, LGPL, EPL)
  - Currently, permissive licensing is trending (MIT, BSD, Apache)
- Several flavours of copyleft (weak, strong, cloud)
  - Requires more or less downstream freedom
  - Impacts compatibility and commerciability
- Not to be confused with open access / available source
  - Much so-called OSS is not… comes with commercial, sectoral or other restrictions

Choosing licenses – Key issues

- Technological
  - What are your direct dependencies?
  - What software does yours interact with?
  - What is your architecture (cloud…)
- Legal
  - Permissive or copyleft? (and what degree of copyleft?)
  - What impact with the license have on downstream uptake and use?
- Community / business
  - What does your community / ecosystem use?
  - How commercially oriented is the software?
  - What are the alternatives / competition?
  - Does it align with your potential business model (where is the value)

Managing licensing

- Inbound and outbound licenses
- Compatibility
- SBOMS and equivalent
- Linking / software interactions, architecture
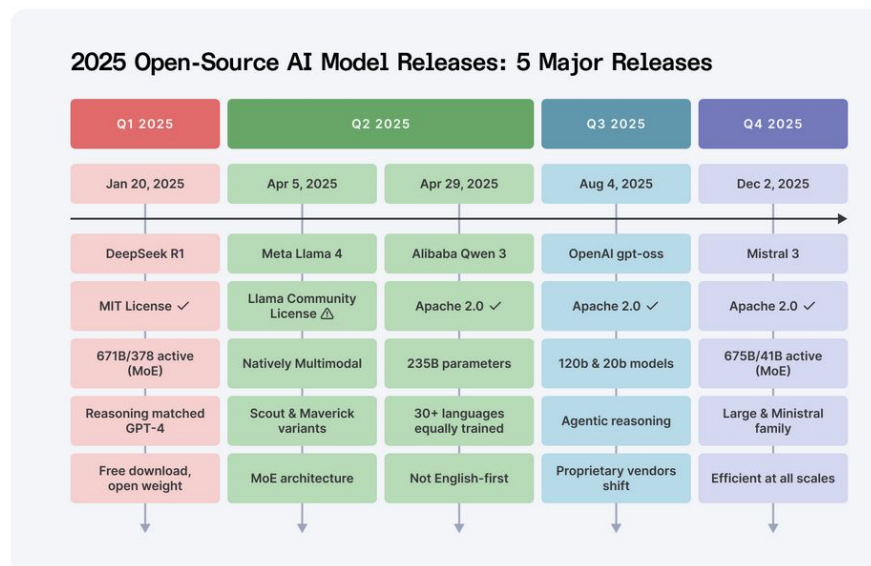- Repository transparency

Common pitfalls

- No license
- No SBOM (inventory of components)
- Incompatible dependencies
- "Available source" (open access)
- No alignment with community, partners, business model

Open source AI (licensing)

Open Source AI is BIG and TRENDING

Quick Reference for Model Selection

Model	Total Params	Active Params	License	Best For	Deployment	Key Trade-off
DeepSeek R1	671B	37B	MIT	Reasoning, Math, Code	Local / API	Chinese geopolitical concerns
Llama 4	405B+	-	Llama Comm.	General + Vision	Your Cloud	EU licensing restrictions
Qwen 3	235B	-	Apache 2.0	Multilingual, Global	Anywhere	Less Western dev community
Mistral Large 3	675B	41B	Apache 2.0	Frontier Reasoning	Any Scale	Slightly slower inference
OpenAI gpt-oss	120B / 20B	-	Apache 2.0	Agentic, Tool-use	Production	API cannibalization risk

2025 Open-Source AI Model Releases: 5 Major Releases

Q1 2025	Q2 2025		Q3 2025	Q4 2025
Jan 20, 2025	Apr 5, 2025	Apr 29, 2025	Aug 4, 2025	Dec 2, 2025
DeepSeek R1	Meta Llama 4	Alibaba Qwen 3	OpenAI gpt-oss	Mistral 3
MIT License ✓	Llama Community License ⚠	Apache 2.0 ✓	Apache 2.0 ✓	Apache 2.0 ✓
671B/378 active (MoE)	Natively Multimodal	235B parameters	120b & 20b models	675B/41B active (MoE)
Reasoning matched GPT-4	Scout & Maverick variants	30+ languages equally trained	Agentic reasoning	Large & Ministral family
Free download, open weight	MoE architecture	Not English-first	Proprietary vendors shift	Efficient at all scales

The 2025 open-source explosion

- Trending topics
  - Open AI technologies
    - Open models (Deepseek, Qwen, Mistral, etc.)
    - Open agentic tools and frameworks (MCP, A2A, LangGraph…)
    - Open Data (Hugging face?)
    - Open RAG (docling)
    - Vector databases
  - Licensing / legal issues
    - Permissive licensing
    - RAIL – responsible AI
    - Contract frameworks
    - Regulatory compliance

Open source AI legal and licensing…

- A nightmare, let's not go there
- AI artefacts / items are NOT the same as pure software
  - Data, models, weights, infrastructures
- Copyright concepts don't work
  - Derivative work of data or models?
  - Copyright protection for models, output?
- Proposal for Open Source AI Definition (OSI)
  - Open source on all code and documentation
  - Data: complete and accurate description, but not open data requirement
  - Specific licenses for AI (creative commons style, with flavours?)
- However… most models are released under a known open source software license (MIT, Apache…)
  - Doesn't account for data (required to "adapt and improve" the AI)
- Regulatory and liability considerations
  - Llama 3.2 onwards: not for EU users
  - AI Act compliance for "putting on the market"

This raises more questions than answers ... for the moment

Soooo... Question time!

Malcolm Bain

Across Legal - Barcelona

mbain@acrosslegal.com