Foundation Model
Due Diligence Toolkit

Type: Selection Checklist

Target Audience: AI Product Managers, Enterprise Architects, Procurement

This toolkit provides a structured approach for evaluating and selecting Large Language Models (LLMs) or Vision models for enterprise use. Complete this assessment before vendor selection.

Model Information

Field	Details
Model Name/Version	
Provider	
Model Type	[ ] LLM [ ] Vision [ ] Multimodal [ ] Other: _______
Deployment Option	[ ] API [ ] Self-hosted [ ] Cloud Instance
Evaluation Date	
Evaluator	

1. Capability Testing

Evaluate performance on specific, representative organizational tasks.

- ☐ Define 10+ representative test cases from actual business scenarios
- ☐ Test accuracy on domain-specific terminology and concepts
- ☐ Evaluate instruction-following capability
- ☐ Test multi-turn conversation coherence (if applicable)
- ☐ Benchmark response latency under expected load
- ☐ Compare results against at least 2 alternative models

2. Limitation Mapping

Document known failure modes and hallucination tendencies.

- ☐ Review provider's model card for documented limitations
- ☐ Test for hallucination on factual questions relevant to your domain
- ☐ Identify topics where model refuses to respond

□ Document context window limitations and their impact
□ Test edge cases and adversarial inputs
□ Assess multilingual capabilities if required

3. Safety Review

Review the provider's red-teaming and safety testing reports.

□ Obtain and review provider's safety/system card
□ Review third-party safety evaluations (if available)
□ Test for harmful content generation
□ Evaluate bias in outputs across demographic groups
□ Assess jailbreak resistance
□ Review content filtering/moderation capabilities

4. Cost Modeling

Project per-token or infrastructure costs for expected production volume.

Cost Factor	Estimate	Notes
Input tokens (per 1M)	$	
Output tokens (per 1M)	$	
Estimated monthly volume		tokens
Monthly API cost	$	
Fine-tuning cost (if applicable)	$	
Infrastructure cost (self-hosted)	$	
Total Year 1 Cost	**$**	

☐ Model costs at 2x and 5x expected volume
☐ Compare cost-per-task across candidate models
☐ Factor in prompt engineering/optimization costs

5. Exit Strategy

Plan for model migration if provider changes terms or discontinues service.

☐ Review contract termination clauses and notice periods
☐ Assess data portability (can you export fine-tuning data?)
☐ Identify alternative models with similar capabilities
☐ Estimate migration effort and cost
☐ Design prompts to be model-agnostic where possible
☐ Document API abstraction layer requirements

6. Vendor Assessment

☐ Verify data handling policy (is customer data used for training?)
☐ Review security certifications (SOC 2, ISO 27001)
☐ Assess SLA guarantees (uptime, latency)
☐ Review IP indemnification terms
☐ Evaluate vendor financial stability and market position
☐ Review model versioning and deprecation policy

Decision Summary

Recommendation: [ ] Approve [ ] Approve with Conditions [ ] Reject

Conditions/Notes: ____

Evaluator Signature: ________________________ **Date:** _____________