



National Intellectual Property Rights Coordination Center and Michigan State University Center for Anti-Counterfeiting and Product Protection

Sample Questions for Potential Artificial Intelligence Partners and Vendors

This document provides a list of representative questions to help guide discussions with potential artificial intelligence (AI) partners and vendors. While non-exhaustive, these questions can serve as a guide to help you assess an AI provider's approach to data curation, AI model training and refinement, data retention and sharing and AI model output.

Data Curation

High quality, diverse training data is crucial to developing accurate AI models. These questions will help you understand an AI vendor's approach to data curation.

- Where does the data that you use to train your AI model come from?
- How do you identify data sources?
- How do you document where you get training data?
- What do you do to ensure that your training data does not include protected intellectual property?

Training and Refinement

Understanding how AI vendors train and update their models can be an important factor to identifying if they are a good fit for your brand's use case.

- How long does it take to train your AI models?
- How often do you update training data?
- At what frequency do you deploy model updates?
- How do you deploy model updates and communicate deployment to users?

Output

It is important to understand how a vendor evaluates its AI model output and ensures high-quality, accurate results.

- How do you ensure that your model does not produce biased outputs?
- Who owns user-directed model outputs (e.g., who owns an AI-generated image created based on a user's text input)?
- Who accepts liability for the output of your AI model?
- By what standard do you evaluate output and monitor performance?

Data Retention and Sharing

How a vendor safeguards your data is an important factor to consider when choosing an Al partner.

- Do you store user input?
- How long do you store user input?
- Is stored user input used to retrain or refine your Al model?
- Who has access to stored user input?
- How do you vet these individuals?