Questions to Ask Vendors About AI

Many vendors providing services to libraries and archives have recently started offering "AI" products. This frenzy to offer AI services doesn't engage with the substantive ethical critiques against using AI, and there are serious possibilities that adopting AI tools may undermine library operations. These questions can help you critically engage when encountering someone trying to market you an "AI product."

---

1. Start here: Basic and technical functions of the tool

- What type of "AI" tool is this, and what aspect of my work does this problem address?
- How is this product different from other pieces of your product suite?
- Did your company build this tool?
- If this tool uses an existing API, what is the value add?
- What tech support do we get with this?
- How is this product licensed? What is the pricing model?
- How long has this product been on the market, how long has this product been under development, and what environments has it been tested within?
- Who owns this product in your organization (e.g. Engineering, Product, Marketing)?
- What's your quality control procedure for this product? How is this tracked and reconciled in your model? (e.g. how does your product handle the following common mechanical errors?)
  - Rather than saying "I don't know", instead providing falsified answers when answers to a query are null/nonexistent
  - Providing answers that confirm the assumed biases of the user in order to establish rapport
  - Ingesting poison-pill data that corrupts that tool and reduces its accuracy rate when answering queries
- Do you have data to show the tool's accuracy over time?

2. Environmental considerations

- Are you familiar with the <u>ALA Core Value of Sustainability</u>? How does the energy usage of this product integrate with libraries' commitment to curtail our resource consumption?
- What are your plans for projected growth in terms of energy and data usage?
- Can you explain in plain language the energy costs per query?
- Who from the organization is responsible for sustainability? Where do you publish ESG (environmental social governance) information? Is this different in the U.S. than it is for other North American, European, or world-wide consumers?

3. Labor expectations

- How does your company foresee its own use of generative AI? Has it replaced any of its workforce with AI?
- What evidence can you share about how this tool will make my work more efficient without adding more to my workflow (learning new tools and troubleshooting)?
- For training and prompting, what kinds of content were/are human data workers enlisted to reinforce and under what conditions did/do they labor?

4. Copyright, data, and privacy considerations

- The <u>ALA Bill of Rights</u> protects patrons' right to privacy and confidentiality in their library use.
  - What are the privacy and confidentiality guardrails for patrons?
  - What patron data is collected? How is that information stored? With whom is it shared? Can you share a privacy policy that specifies this AI tool?
- The <u>ALA Code of Ethics</u> states: We respect intellectual property rights and advocate balance between the interests of information users and rights holders.
  - Can you be certain whether your product has been trained, is being trained, or will be trained on data in such a way that it might violate copyright laws in the US or other countries where the product is being licensed to customers?
- What are the citation practices for output generated by your product?
- What are copyright expectations for output generated by your product?
- Where do you source the data that you use as training material for your product?

5. Revenue expectations

- What percentage of your company's revenue is budgeted to come from AI?
- How much capital has been invested in this?

Glossary of terms

AI: shorthand for "artificial intelligence." Often functions as an umbrella term for loosely connected technologies, and is increasingly used to market computational products as "more" than they are or as "intelligent" in a way that resembles or imitates human consciousness.

AGI: shorthand for "artificial general intelligence," or a type of AI that is "beyond" human comprehension, or is "godlike" in some way. It may be brought up by marketers and advertisers seeking to manufacture the idea that AI products are not only here to stay but are inherently moving in the direction of greater complexity and competence.

AI Hallucination: an AI-generated response that presents false information as fact. One of many anthropomorphizing terms used in descriptions of AI; more accurately can be referred to as a "systems error."

Algorithm: a sequence of mathematical instructions, typically used to solve problems or perform a computation.

Bot: (chatbot, spambot, social bot) a computer program that is trained to respond to queries or prompts with natural language.

Generative AI: an algorithmic tool that typically relies on immense datasets used for training the tool, which then operates by churning that data to "generate" predictive responses to queries. Other terms you might come across are:

    LLM: "large-language model" and **SLM:** "small-language model"

Generative Pretrained Transformer (GPT): a machine learning architecture that requires immense computational power and massive data sets in order to provide seemingly intelligent responses to user prompts.