Teacher Answer Key

Case Study

1. The strongest open source arguments for Option B are: local deployment (patient data never leaves the clinic's own servers, which is critical for medical privacy), democratization of access (the clinic cannot afford $2,400/month), and independent inspection (their IT staff can examine the model's behavior rather than trusting an external company's safety claims).

2. Option B risks: less powerful models may produce less accurate medical summaries, potentially affecting patient care. With no corporate support, errors or failures must be handled by a volunteer IT person. Community-maintained open source models may have inconsistent quality and no accountability structure if the AI gives harmful advice.

Multiple Choice

1. D

2. B

3. C

Matching

1-B, 2-C, 3-E, 4-A, 5-D

Fill-in-the-Blank

1) Llama, 2) millions, 3) vulnerabilities

Short Answer (Sample Responses)

Q1: Open source enables safety research by letting anyone inspect a model for problems — many safety issues have been found by independent researchers who couldn't work with closed models. But open source also enables safety removal: once weights are public, bad actors can strip out safety measures and use the model to generate harmful content. The same transparency that enables protective research also enables malicious modification.

Q2: Closed source frontier AI is expensive — enterprise API costs can reach thousands of dollars monthly, restricting access to well-funded companies. Open source models allow nonprofits, small businesses, researchers in developing countries, and individuals to access comparable AI capabilities for free. This matters for equity because AI access shapes economic opportunity, and

concentration of AI capability in wealthy companies concentrates the benefits of the technology.

Q3: Sample — Against mandatory release: frontier models represent enormous investment, and mandatory release would eliminate the business model funding their development. It would also allow any actor worldwide to remove safety measures, creating risks impossible to recall. For mandatory release: public models enable independent safety scrutiny impossible with closed systems. A middle position — mandatory disclosure to government safety reviewers without full public release — might balance both concerns.