**Bài tập nhóm 01**

**Câu 1:**

Một trong những nguyên tắc đạo đức quan trọng là tính minh bạch. Điều này giúp cho các nhà phát triển AI đảm bảo rằng quá trình phát triển, triển khai và sử dụng AI được thực hiện một các rõ ràng và minh bạch, bao gồm cả quá trình huấn luyện và thu thập dữ liệu. Tính minh bạch cũng giúp xây dựng lòng tin từ người dùng, cho phép họ hiểu rõ cách mà AI đang hoạt động và tại sao nó đưa ra quyết định như vậy. Vì vậy, tính minh bạch là một yếu tố quan trọng trong việc phát triển AI đáng tin vậy và giúp tăng cường lòng tin của người dùng vào AI. Nó cũng giúp đảm bảo rằng AI được sử dụng một cách công bằng và đúng đắn, giúp tạo ra một tương lai có trách nhiệm và bền vũng trong lĩnh vực trí tuệ nhân tạo.

**Câu 2:**

Một ví dụ về việc sử dụng trí tuệ nhân tạo không đạo đức là việc sử dụng AI để tạo ra các hình ảnh, video giả và file âm thanh mạo danh người thân để lừa đảo và gây hại cho người khác.

Ví dụ, một ai đó có thể sử dụng AI để tạo ra một video giả mạo một nhân vật nổi tiếng nói ra những điều mà học chưa từng nói ra, nhằm tạo tin đồn giả mạo và gây ảnh hưởng đến danh tiếng của họ. Hay là kẻ xầu sử dụng nó tạo ra những video âm thanh báo cho gia đình là có em nhỏ bị tai nạn và hối thúc người nhà chuyển tiền viện phí phẫu thuật ngay lập tức. Chúng lợi dụng vào sự dễ tin người và lo lắng của mọi người trong thời điểm đó khi nghe tin không may như vậy.

Điều này cũng đã cho thấy rõ ràng tại sao chúng ta cần có quy định đạo đức trong phát triển trí tuệ nhân tọa. Nếu không có, chúng ta có thể đối diện với các tác động tiêu cực và nguy hiểm đến lòng tin, danh tiếng và an toàn của mọi người. Chính vì thế, các quuy định đạo đức trong phát triển trí tuệ nhân tạo là rất cần thiết để đảm bảo rằng AI được sử dụng một cách đúng đắn và mang lại lợi ích tốt đẹp cho cộng đồng.

**Câu 3:**

Việc sử dụng AI trong bảo mật thông tin có thể gây ra nhiều rủi ro như ra quyết định thiên vị, thiếu khả năng giải thích và minh bạch, tiềm năng bị lạm dụng hoặc lạm dụng, dễ bị tấn công, thiếu giám sát của con người, chi phí cao và mối quan tâm về quyền riêng tư.

Để giảm thiểu những rủi ro này, các quy định như CPRA và NAIIA được đề ra. CPRA cung cấp thêm các biện pháp bảo vệ quyền riêng tư cho người tiêu dùng và yêu cầu doanh nghiệp công bố các thực tiễn thu thập và chia sẻ dữ liệu của họ. Còn NAIIA nhằm thúc đẩy nghiên cứu và phát triển AI trong khi đảm bảo rằng nó được phát triển theo cách an toàn, bảo mật và có ích cho xã hội. Từ đó thấy được những quy định này vô cùng quan trọng để giảm thiểu các rủi ro liên quan đến việc sử dụng AI trong bảo mật thông tin.

**Bài tập nhóm 02**

**Câu 1:**

Trong trường hợp này, công ty có thể đã vi phạm đạo đức AI. Việc loại bỏ những ứng viên có khả năng bị trầm cảm hoặc mang thai trong tương lai có thể coi là hành vi phân biệt đối xử và vi phạm quyền bình đẳng cơ hội trong việc làm. Điều này có thể dẫn đến những hậu quả tiêu cực cho các ứng viên bị loại bỏ và gây ra những tranh cãi về đạo đức trong cộng đồng. Ngoài ra, việc sử dụng AI để ra quyết định trong tuyển dụng cũng có thể gây ra những vấn đề về minh bạch và khả năng giải thích, khiến cho ứng viên sẽ thấy khó hiểu lý do tại sao một ứng viên đó bị loại bỏ. Điều này có thể gây ra những tranh cãi và mất lòng tin vào hệ thống tuyển dụng của công ty.

Tóm lại, việc sử dụng AI để loại bỏ những ứng viên có khả năng bị trầm cảm hoặc mang thai trong tương lai có thể vi phạm đạo đức AI và gây ra những hậu quả tiêu cực cho các ứng viên và công ty trong thời điểm hiện tại và trong tương lai. Vì vậy, công ty cần cân nhắc lại quyết định của mình và xem xét các tiêu chí tuyển dụng khác, có liên quan đến năng lực và kinh nghiệm để đảm bảo tính công bằng và đúng đạo đức trong việc tuyển dụng.

**Câu 2:**

Trung Quốc đã đặt mục tiêu trở thành bà chủ toàn cầu về trí tuệ nhân tạo với giá trị ngành công nghiệp lên đến 150 tỷ đô vào năm 2030. Kế hoạch này bao gồm các sáng kiến và mục tiêu cho việc nghiên cứu và phát triển, công nghiệp hóa, phát tiển nhân tài đất nước, giáo dục và thu thập kỹ năng, thiết lập tiêu chuẩn và quy định, chuẩn mực đạo đức và an ninh trong lĩnh vực trí tuệ nhân tạo. Kế hoạch hành động đổi mới AI cho các cơ sở giáo dục đại học đã đưa ra một số mục tiêu nhằm tăng cường đang kể lực lượng nhân tài AI của Trung Quốc và chương trình giảng dạy AI của các trường Đại học vào năm 2030.

**Câu 3:**

Trong các luật về AI được áp dụng ở Trung Quốc từ năm 2020 đến nay, bộ luật đáng chú ý đó là bộ luật về bảo vệ thông tin cá nhân (Personal Information Protection Law). Bộ luật rất hữu ích vì nó bảo vệ những dữ liệu quan trọng của mọi người và giúp đảm bảo sự cân bằng giữa sự phát triển và bảo vệ quyền riêng tư.

Việt Nam có thể học hỏi bộ luật trên để xây dựng một khung pháp lý cho việc quản lý dữ liệu đại chúng tương tự. Việc áp dụng sẽ hạn chế việc lạm dụng thông tin cá nhân làm những việc xấu và khuyến khích các tổ chức sử dụng dữ liệu một cách hợp lý và minh bạch.

Tuy nhiên, để phù hợp với đời sống và tình hình ở Việt Nam, bộ luật này có thể được cải tiến để cho phù hợp. Ví dụ, Việt nam có thể tăng cường việc đảm bảo tính công khai và minh bạch trong việc sử dụng dữ liệu. Ngoài ra, cần có sự hỗ trợ cho các doanh nghiệp để hiểu rõ những yêu cầu và quy định của bộ luật này nhằm để thực hiện các hoạt động của mình một cách hợp lú và đúng luật hiện thời đề ra.