

シミュレーション演習：

安全・危険なAIプロントを見分ける

概要

AIツールは業務の効率化に大変役立ちますが、一歩使い方を誤ると重大なセキュリティリスクを生み出します。職場が指定した承認済みのAIツールは、一般的な無料のAIに比べてデータ保護が厳重です。しかし、どのような情報を入力（プロンプト指示）するかという利用者の扱い方次第では、深刻な脆弱性を作ってしまうかねません。だからこそ、AIに入力するプロンプトの安全性を常に意識する習慣が不可欠です。

たとえば、不適切なプロンプトを入力したせいでAIが誤った回答を出力し、誤認を招いてしまったり、本来は取り扱いに注意が必要な社内の重要情報が流出してしまったりする危険性があります。このアクティビティでは、社内で認められた公式AIツールを使って業務効率を高めつつ、同時に安全なプロンプト作成のベストプラクティスを身につけることを目指します。

役割

進行役：

演習全体の進行ペースを管理し、ディスカッションをリードしながら、参加者全員が気軽に発言できるよう場をまとめます。

参加者：

演習のルールに沿って自分の考えを共有し、提示された事例についてグループで話し合います。

学習内容

進行役は、参加者に次のシナリオを説明してください。【シナリオ】職場で全従業員が利用できる公式のAIツールが導入されました。このツールは社内専用の安全な環境で運用されるため、入力したデータが外部に公開されたり、一般のAIの学習データとして勝手に使われたりすることはありません。ただし、環境が安全だからといって、何を入力しても良いわけではありません。プロンプトを作成する際は、必ず次のセキュリティ規定を守る必要があります。

- 【厳禁】個人情報の入力（氏名、社員ID、生年月日などの個人特定情報、および健康・医療データなど）
- 【厳禁】機密情報の入力（企業の営業秘密や社外秘に指定されている文書など）
- 【厳禁】認証情報の入力（銀行の口座情報、各種ログインID、パスワードなど）

導入から1か月が経ち、上司から社内でのプロンプト利用状況に関するレポートが共有されました。みなさんでレポートに載っているプロンプトを確認し、「安全なもの」と「危険なもの」に仕分けてください。

パート1：手がかりを見つける

🕒 所要時間：15分

【準備するもの】

- ・ タイマーまたはストップウォッチ
- ・ シナリオの配布資料
- ・ 筆記用具
- ・ メモ用紙

【アクティビティの流れ】

- ・ 参加者を数人のグループに分けます。
- ・ 各グループにワークシート（プロンプトの一覧資料）と筆記用具を配ります。
- ・ グループで資料に書かれたプロンプトを確認します。
- ・ 安全だと思うプロンプトには下線を引いてもらいます。（例：具体的なプロジェクト名や顧客名を出さず、抽象的な表現に置き換えているなど）
- ・ 危険だと思うプロンプトには丸をつけてもらいます。（例：機密情報や個人情報をそのままAIに入力しているなど）
- ・ 職場のルールを守り、承認された安全なデータだけを取り扱うことの重要性を全体に共有します。
- ・ 次の2つのテーマについて、グループで意見をまとめてもらいます。①危険だと判断したプロンプトは、どう書き換えれば安全に使えたと思いますか？②安全だと判断したプロンプトのなかで、自身の業務に役立ちそうなものはありましたか？
- ・ 制限時間は15分です。グループで話し合いながら、メモを取ってもらいます。

パート2：ディスカッション

🕒 所要時間：10分

【振り返りの質問】

- ・ プロンプトのどのような点を見て「これは危険だ」と気づけましたか？
- ・ 最初は安全そうに見えて、実は危険だったプロンプトはありましたか？
- ・ AIを安全に利用する方法の中で、すでに知っていたものはありましたか？
- ・ 普段の業務やプライベートで、AIを利用する際に気をつけている点や工夫している点があれば教えてください。

【想定される回答・意見の例】

- ・ 日常の業務の補助としてAIを使う場合、機密情報が含まれていなければ問題ない。職場が承認した社内専用のAIツールを利用しているので、これらのプロンプトは安全だ。入力したデータが外部の公開AIモデルの学習に使われることはない。
- ・ 危険なサイン：アカウントの安全性を脅かす行為、業務ツールへの個人のプライベートデータの混入、および企業の知的財産や機密情報の開示につながる要求。

まとめ

🕒 所要時間：5分

今回の演習を振り返り、日々のセキュリティ意識へとつなげます。危険なプロンプトに気づくことは、あくまで防御の第一歩にすぎません。本当に重要なのは、組織の規定に従って、適切な表現に修正した上でAIを活用することであると伝えます。

業務の効率アップを図りつつ、安全にAIを使いこなすためのプロンプトの作成方法について、参加者からの質問を受け付けます。

職場でのAIの使い方や工夫している点などについて、参加者同士で話し合いましょう。

ID	従業員のAIプロンプト	カテゴリ
A	荷物の追跡サービスをリニューアルするから、目を引くキャッチコピーを10個作って。	安全 ：クリエイティブなアイデア出しやブレインストーミングにAIを活用しています。
B	これが私の今週のスケジュールね。時間を最も効率よく使えるように、タスクを作業の難易度順に並び替えて整理して。	安全 ：社内専用の安全な環境であるため、自身のタスク管理や業務効率化のために活用できます。
C	上司に報告するために、この極秘の企業合併に関する資料の要点をまとめないといけないの。要約文を作成して。	危険 ：いくら社内専用の安全なツールであっても、企業の経営に関わるような最重要機密文書は、別途厳しい取扱規定が設けられていることがほとんどです。今回の演習のルールと照らし合わせると、AIツールにアップロードする行為は規定違反にあたります。
D	アジアに新設した倉庫のサービスに満足してくれている取引先宛てに、お礼のメールを作成して。	安全 ：特定の顧客名や具体的な個人情報を出さず、一般的なビジネスメールの作成を依頼しています。
E	いまゲストアカウントでこのAIツールを使ってるんだけど、忘れないように、私の銀行口座のパスワードをここに記録しておいてもらってもいい？	危険 ：たとえ職場が承認している安全なAIツールであっても、パスワードやログイン情報といった認証情報を入力することは絶対に避けてください。
F	締め切りが昨日だったタスクがあるんだけど、同僚にその進捗状況を確認したいからメールを作ってもらっていい？関係を壊したくないから、気まづくならないような丁寧な文面にして。	安全 ：日常業務におけるちょっとしたビジネス文章の作成や表現の改善にAIを活用しています。
G	これがうちの取引先の一覧。「氏名・誕生日・所在地・取引開始日」が書いてあるから、これをもとに一人ひとりに向けたオリジナリティのあるバースデーメッセージを作成して。	危険 ：氏名や生年月日といった個人を特定できる情報をAIにそのまま入力するのは明確な規定違反です。プロンプトを作成する際は、顧客や取引先の個人情報を絶対に含めないでください。
H	IT業界の職種でよく使われる「プロジェクトマネージャー」と「プロダクトマネージャー」の根本的な違いについてわかりやすく教えて。	安全 ：自社の機密情報や個人情報を一切含まない、一般的な知識や概念、公開情報についての質問です。

【配布資料】ワークシート

A～Hのプロンプトは、従業員が実際に社内のAIに入力した内容です。職場のセキュリティ規定を思い出しながら、これらを「安全なもの」と「危険なもの」に分類し、危険なプロンプトはどのように書き直すべきかグループで話し合ってください。

ID	従業員のAIプロンプト	カテゴリ
A	荷物の追跡サービスをリニューアルするから、目を引くキャッチコピーを10個作って。	
B	これが私の今週のスケジュールね。時間を最も効率よく使えるように、タスクを作業の難易度順に並び替えて整理して。	
C	上司に報告するために、この極秘の企業合併に関する資料の要点をまとめないといけないの。要約文を作成して。	
D	アジアに新設した倉庫のサービスに満足してくれている取引先宛てに、お礼のメールを作成して。	
E	いまゲストアカウントでこのAIツールを使ってるんだけど、忘れないように、私の銀行口座のパスワードをここに記録しておいてもらってもいい？	
F	締め切りが昨日だったタスクがあるんだけど、同僚にその進捗状況を確認したいからメールを作ってもらっていい？関係を壊したくないから、気まづくならないような丁寧な文面にして。	
G	これがうちの取引先の一覧。「氏名・誕生日・所在地・取引開始日」が書いてあるから、これをもとに一人ひとりに向けたオリジナルティのあるバースデーメッセージを作成して。	
H	IT業界の職種でよく使われる「プロジェクトマネージャー」と「プロダクトマネージャー」の根本的な違いについてわかりやすく教えて。	