

AIを安全に使うために 知っておきたい危険な落とし穴

安心して、堂々とAIを使い続けるために

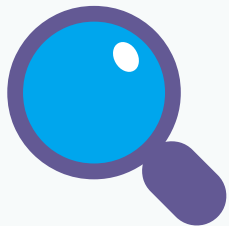


本日のゴール

3つのステップで、AIとの付き合い方がちょっと変わります。

01

気づく



よくある「うっかり」に
気づけるようになる

02

わかる



なぜ危険なのか、
仕組みをざっくり知る

03

安心する



対策を知って、
堂々と使い続けられる



今日はこの3ステップで「知らなかった」を「知ってる」に変えます。



本日の流れ

全部で9つのリスクを、対策とセットで見えていきます。

1



なぜ危険なのか

AIを使う前に知っておきたい心構えの話

2



パソコン・データ

うっかり本番環境を壊してしまうリスク

3



情報漏洩

APIキーや個人情報が漏れてしまうリスク

4



著作権

作った作品の権利をめぐる注意点

5



肖像権

実在の人物の写真を元にした生成に関する注意点

6



損害賠償・信用

確認不足がトラブルにつながるリスク

7



だまされるリスク

AIの間違いや、なりすまし詐欺への備え

8



プロンプトインジェクション

仕込まれた指示に気づかず情報が漏れるリスク

9



検算・シャドーAI

鵜呑みにする・こっそり使うことのリスク



なぜ危険と隣り合わせなのか

AIは便利な道具。でも、道具には「正しい使い方」があります。



ブレーキやハンドル操作、落ち着いた対処ができないのに
車に乗るような行為なのかなと思っています。

AIも同じです。「何をしているか」を知らないまま使うと、思わぬ事故につながる可能性があります。
でも大丈夫。危険なポイントは、実はそんなに多くありません。これから一つずつ、
「知れば防げる」落とし穴を、対策とセットで見ていきましょう。



難しく考えなくて大丈夫。今日知れば、もう安心です。



パソコン・データを壊すリスク

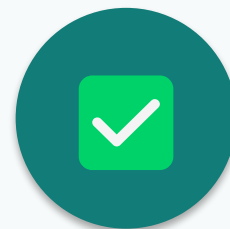
「あとはよろしく」が、思わぬ事故につながる可能性があります。



ありがちな失敗

誰にでも起こりうること

- ・ 良かれと思って、社内システムの整理・自動化をAIコーディングツールに丸ごとお任せし、確認せずに実行を承認し続けた
- ・ その結果、必要なデータを誤って削除。復旧費用を勤務先から請求されてしまうケースも



できる対策

今日からできる工夫

- ・ 大事な操作の前は「何をするか」を一度説明させてから実行する
- ・ 消えたら困るデータは、作業前にバックアップを取っておく
- ・ 慣れるまでは、詳しい人と一緒に本番環境を触る



情報漏洩のリスク

「うっかり公開」が、一番身近な落とし穴です。



ありがちな失敗

誰にでも起こりうること

- ・ 良かれと思って、顧客リストや取引先情報をAIに読み込ませて「整理しといて」と頼んだ
- ・ そのデータが外部に漏れ、勤務先が取引先から損害賠償を求められ、担当した自分も責任を問われることに



できる対策

今日からできる工夫

- ・ 個人情報・機密情報はAIに入力しない
- ・ どうしても必要なときは、会社のルール・許可を確認してから使う
- ・ 不安なときは一人で判断せず、まわりに相談する



著作権のリスク

「AIが作ったから大丈夫」とは限りません。



ありがちな失敗

誰にでも起こりうること

- ・ 良かれと思って、会社のチラシやSNS投稿にAIが作ったイラストをそのまま使用
- ・ 既存の作品に似ていたことが後から分かり、著作権侵害として勤務先に損害賠償請求が届き、担当した自分も厳しく問われることに



できる対策

今日からできる工夫

- ・ 商用で使う前に、そのAIツールの利用規約を確認する
- ・ 既存の作品にそっくりでないか、自分の目でも一度確認する
- ・ 不安なときは公開前に周りに相談する



肖像権のリスク

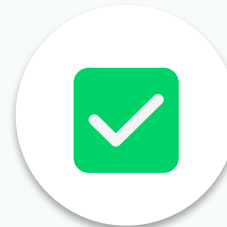
写真を元にした画像生成には、写っている「人」への配慮が必要です。



ありがちな失敗

誰にでも起こりうること

- ・ 良かれと思って、AIラボで作ったオリジナルLINEスタンプの元絵に、参加者の顔写真をそのまま使用
- ・ 本人は同意していても、隣に写り込んだ友だちの同意までは取れておらず、後から指摘されるケースも



できる対策

今日からできる工夫

- ・ 写真を元に生成物を作る前に、写っている全員の同意を確認する
- ・ 特に子どもが写る写真は、保護者・本人の同意を事前に取り
- ・ SNS等に公開する前は「顔が分かる生成物かどうか」を最終チェックする



「本人の同意」は、写真に写っている全員分。ここまで意識できると安心です。



損害賠償・信用のリスク

「確認せずそのまま送信」が、一番の落とし穴です。



ありがちな失敗

誰にでも起こりうること

- ・ 良かれと思って、取引先へのお詫びメールをAIに書いてもらい、内容を確認せずそのまま送信
- ・ 実際とは違う約束をしてしまい、勤務先がその対応に追われ、担当者個人としても信用を落とすことに



できる対策

今日からできる工夫

- ・ 送信・公開の前に、必ず人の目で読んで確認する
- ・ 事実と違う点や失礼な表現がないかをチェックする
- ・ 「最終判断は人が持つ」を合言葉にする



だまされるリスク

AIの「自信満々な間違い」と、悪用の手口に注意。



ありがちな失敗

誰にでも起こりうること

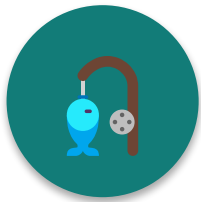
- ・良かれと思って、AIが自信満々に教えてくれた情報をそのままお客様に案内。実は誤った情報で、勤務先がクレーム対応に追われ、案内した本人も信用を落とすことに
- ・AIで合成された「上司の声」で送金を指示され、確認せずに振り込んでしまうケースも



できる対策

今日からできる工夫

- ・大事な情報はAIの回答だけで判断せず、必ず別の情報源でも確認する
- ・お金や重要な指示は、電話やメッセージだけでなく、別の方法でも本人確認する



プロンプトインジェクションのリスク

メールや資料に仕込まれた「指示」に、AIが従ってしまうことがあります。



ありがちな失敗

誰にでも起こりうること

- ・良かれと思って、取引先から届いたメールをそのままAIに読み込ませ「返信案を作って」と頼んだ
- ・メール内に仕込まれた指示にAIが従い、社内の別情報を外部に送る動きをしてしまった



できる対策

今日からできる工夫

- ・知らない相手のメール・添付・Web情報をAIに読ませる前に、不審な指示が紛れていないか目を通す
- ・AIの「学習に使わない」設定（オプトアウト）を標準にし、入力情報が外部の学習に使われないようにする
- ・機密性の高いやり取りは、AI連携ツールに丸ごと接続せず必要な範囲だけを渡す



検算不足・シャドーAIのリスク

「AIが言ったから」を鵜呑みにすることと、「隠れて使う」ことの2つの落とし穴。



ありがちな失敗

誰にでも起こりうること

- ・ 良かれと思って、AIが出した数字や分析結果をそのまま会議資料に転記。実は根拠のない数字（ハルシネーション）で、後から指摘され信用を落とした
- ・ 会社が正式導入していないからと、個人のスマホアプリでこっそりAI（シャドーAI）を使い、社外秘の情報を入力してしまっていた



できる対策

今日からできる工夫

- ・ AIが出した数字・固有名詞・引用は、必ず元データや別の情報源で検算してから使う
 - ・ 「勝手に使う」より「正式に申請して使う」ほうが記録も残り安心。公式導入をむしろ後押しする
 - ・ 不安なときは情シス・上長に「これって使っていい？」と一声かける文化にする



安全に使うための鉄則

ここまでの対策を、5つにぎゅっとまとめました。

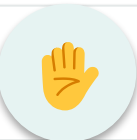
1



秘密は書かない

パスワードやAPIキー、個人情報は入力しない

2



一度立ち止まる

大事な操作の前は、内容を確認してから実行する

3



人が必ず確認

AIの出力はそのまま使わず、人の目でチェックする

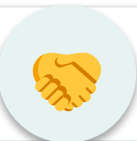
4



規約を確認

権利関係・利用規約を見てから使う

5



一人で抱えない

不安なときは、まわりに相談する



この5つさえ押さえておけば、もう安心してAIを使えます。



ワーク実践（5分）

さあ、ここからは皆さんの番です！



あなたが使っているAIツールで、
「あ、危なかったかも」と思ったことはありますか？

隣の人と2分間シェアしてみましょう。

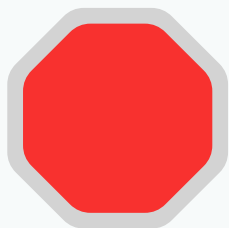


今日から意識する3つのこと

難しいことは忘れても、これだけ覚えて帰ってください。

01

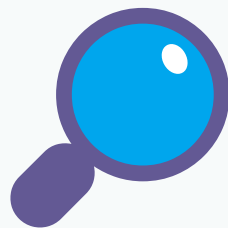
立ち止まる



大事な場面では
一呼吸置く

02

確かめる



AIの答えを鵜呑みにせず
確認する

03

相談する



不安なときは
一人で抱えない



これで今日から、安心して堂々とAIを使い続けられます。

ご清聴ありがとうございました

質疑応答