

FLORA REPORT | AI時代の人事データ シリーズ 01

人事AIの失敗パターン

～判断より前に壊れているもの。

同じデータから違う答えが出る仕組みを、計算で示す～

CONTENTS

目次

—	要点：5つの結論	3	05	出典：説明できない数字は、直すこともできない	8
01	壊れているのは、判断の手前である	4	06	範囲：生成物は、読む人の権限を超える	9
02	比較集団：同じ「上位10%」が3通りある	5	07	公開された失敗事例から分かること	10
03	比較集団：性別でも同じことが起きる	6	08	導入前の点検12項目	11
04	定義：同じ指標名で、別の数字	7	—	本資料の作成方法と出所	12
			—	Floraのご支援：CapLead	13
			—	お問い合わせ	14

この資料の使い方 02から03は、AIの精度とは関係なく結果が入れ替わる仕組みを、計算例で示したものです。自社の名簿で同じ計算を再現できます。04から06は、その計算の手前にあるデータ側の条件です。08の点検表は、AIの導入を検討する段階でそのまま使えます。**本資料の数値は、架空のグループ「サンライズHD」（12社・従業員8,420名）のサンプルデータです。** 実在の企業のデータではありません。

EXECUTIVE SUMMARY

人事AIの失敗は、モデルの精度ではなく、 その手前のデータの状態で決まっている。

1

比較する集団を変えるだけで、選ばれる人が入れ替わる

同じデータ、同じスコア、同じ「上位10%」でも、順位をつける集団を全社にするか会社ごとにするかで結果が変わる。サンプルデータでは一方の会社の取り分が47%増えた。モデルの不具合ではない。算数の結果である。

2

この誤りは、精度検証では見つからない

スコアの付け方も、順位のつけ方も、上位10%を取る処理も正しい。どこにも不具合がないため、検証は通る。比較集団の指定は設定項目として存在しないことが多く、既定値のまま使われる。

3

定義が固まっていない指標は、AIに渡した時点で意味を失う

管理職の範囲や分母の取り方が届け先ごとに違う指標をそのまま読ませると、生成された文章は正しい計算から誤った説明を作る。定義は数字に付いていなければならない。

4

出典がなければ、誤りを直すこともできない

出力が疑われたとき、元データの行まで遡れなければ、正しいのか誤っているのかを確かめられない。撤回もできず、再発防止も書けない。説明責任は精度の問題ではない。

5

生成物は、読む人に見せてよい範囲を自動では守らない

集計値が少人数で個人に還元される、権限のない読者に他社の数字が渡る。生成の段階で範囲を適用しない限り、出力は元データの権限設定を引き継がない。

対象読者 人事・経営企画でAIの導入を検討している方、人事データを扱う情報システム部門、AIの利用に関する社内ルールを作る立場の方。

01 全体像

失敗が起きるのは判断の段階ではない。 データが判断に届くまでの4か所である

人事AIの失敗として語られるのは、たいていモデルの精度か、ブラックボックスであることだ。しかし実際に結果がおかしくなる原因は、その手前にあることが多い。問題は、何を学習したかより前にある。誰と誰を比べたか。どの定義で数えたか。その数字がどこから来たか。誰に見せてよいか。この4つである。

図表1 判断の手前にある4つの壊れ方

① 比較集団

誰と比べて上位なのか指定されていない。既定値のまま全社1本の順位が使われ、等級構成の違いが実力の違いに見える。

→ 02・03

② 定義

管理職の範囲、分母、対象期間が届け先ごとに違う。同じ指標名の数字が社内に複数あり、AIはそのどれかを読む。

→ 04

③ 出典

出力された数字が、どのファイルのどの行から出たかを辿れない。疑われたときに確かめられず、直せない。

→ 05

④ 見せてよい範囲

生成物が、読む人の権限や少人数非表示のルールを超える。元データの制御は、出力には自動では引き継がれない。

→ 06

注：いずれもモデルの性能とは独立した問題で、より高性能なモデルに置き換えても解消しない。出所：Flora株式会社

共通しているのは、**どれも出力を見ただけでは気づけない**という性質である。出てくる文章は自然で、計算は正しく、順位表も整っている。誤りがあるのは数字の中ではない。数字が作られた前提の中にある。

02 比較集団

同じ2,820名から「上位10%」を選ぶ。 順位をつける集団を変えるだけで、片方の会社の取り分が47%増える

図表2 2社の等級構成（サンプルデータ）

会社	従業員	上位等級 (係長級以上)	一般等級
サンライズ商事	1,840名	644名 (35.0%)	1,196名
サンライズ物流	980名	98名 (10.0%)	882名
2社計	2,820名	742名	2,078名

注：上位等級の比率が3.5倍違う。等級が上がるほどスコアが高く出る仕組みは、人事評価でもAIのスコアリングでも共通である。前提：等級の中での実力の分布は2社で同じだと置いている。出所：サンライズHD サンプルデータ（架空のグループ）

図表3 同じ282名を選ぶ3つの方法と、その結果

選び方	サンライズ商事	サンライズ物流
全体で上位10% (多くの製品の既定値)	215名 自社の11.7%	67名 自社の6.8%
会社ごとに上位10%	184名	98名
等級ごとに上位10%	184名	98名
既定値との差	-31名 (-14%)	+31名 (+47%)

注：どの方法も「上位10%を選ぶ」という同じ指示で、合計は282名で変わらない。差を生んでいるのは実力の違いではない。等級構成である。出所：同上

この計算は自社で再現できる 必要なのは、会社ごとの在籍者数と等級別の人数だけである。スコアの中身を知らなくても、等級構成が違えば全社順位でどちらに偏るかは計算できる。先に計算してから、選び方を決める。順番が逆になると、選ばれた後で説明を作ることになる。

03 比較集団

同じことが性別でも起きる。 「候補が少ない」は、判定の前提に依存している

図表4 上位10%に該当した割合（サンプルデータ）

全体の順位で 判定 8.9% 女性	15.4% 男性	同性内の順位で 判定 12.1% 女性	12.4% 男性
-----------------------------------	--------------------	-------------------------------------	--------------------

注：同じデータ、同じスコア、同じ上位10%。変えたのは順位をつける集団だけである。全体順位では1.7倍の差に見え、同性内順位ではほぼ差がなくなる。出所：サンライズHD サンプルデータ（架空のグループ）

「候補になる女性が少ない」という説明は、どちらの判定を使ったかで成り立ったり成り立たなくなったりする。前提を書かずに結論だけを引き継ぐと、事実として社内を回る。

図表5 どちらが正しいのかは、目的によって決まる

全体の順位	今すぐ全社の基準で任せられる人を選ぶ場面には合う。等級構成や職種構成の偏りをそのまま引き継ぐ
同性内・等級内の順位	偏りの影響を除いて候補を広く見る場面に合う。育成対象の選定や、母集団の厚みを測る用途
決めておくこと	どちらを使うか、なぜそれを使うか、いつ見直すか。この3点が記録されていれば、選から外れた人への説明ができる

注：本資料はどちらか一方を推奨しない。問題は選択そのものではない。記録が残っていないことである。

説明の場面で効いてくる 選から外れた人に理由を聞かれたとき、「全社の順位で上位10%に入らなかった」は答えになるが、「なぜ全社の順位なのか」には答えていない。等級ごとに見れば入っていた人に対しては、なおさらである。比較集団の選択は、結果を決めているのに記録が残らない種類の判断だ。

04 定義

「女性管理職比率」という名前の数字が社内に4つある。
AIはそのどれかを読み、どれを読んだかは出力に残らない

図表6 同じ指標名で数字が変わる主な原因

変わるもの	何が違うか	AIに読ませたときに起きること
管理職の範囲	課長級以上か、課長代理を含むか、社内の呼称（マネージャー）か	前年比の説明が、登用の成果として書かれる。実際には範囲を変えただけ
法人の範囲	提出会社のみか、連結子会社それぞれか、グループ合算か	グループ全体の傾向として書かれた文章が、1社だけの数字を根拠にしている
従業員の範囲	就業人員、常時雇用する労働者、正社員のみ、出向者の扱い	分母の違う数字が並び、比率の差が実態の差として解釈される
対象時点・期間	期末時点か期中平均か。事業年度か計画期間か	時点の違う数字が時系列として並び、増減が読まれる
端数処理	小数第1位以下切り捨てか四捨五入か	目標達成・未達の判定が変わる

注：いずれも実務上どれも誤りではない。届け先ごとに求められる定義が違うために、同じ会社の同じ年度で複数の正しい数字が存在する。制度ごとの違いは「人的資本 開示指標 定義対照表 2026」（Flora）に整理している。出所：Flora株式会社

対策は、数字に定義を付けて動かすこと 指標ごとに定義を文書化し、版を付ける。出力側は数値と一緒に定義の版を表示する。制度改正があれば定義側を更新し、その定義を参照しているすべての出力に反映させる。数字と定義が別々に管理されているかぎり、この整合は人が毎回確認するしかない。

生成AIに固有の難しさ 表計算であれば、分母の違う数字を並べても読む人が気づくことがある。生成された文章は、分母の違いを説明せずに滑らかにつなぐ。**読みやすさが、検証の手がかりを消す。**出力に定義の版を明示することが、この性質への最小限の対処になる。

05 出典

出典がない数字は、誤りを証明することも、正しさを証明することもできない

AIが出した数字が疑われる場面は必ず来る。監査人から、投資家から、取締役会から、あるいは選から外れた従業員から。そのとき問われるのは精度ではない。その数字がどこから来たかである。

図表7 「どこから来たか」に答えるための4要素

1 定義書	算出式、法人と従業員の範囲、対象時点、抽出条件、端数処理、改定履歴
2 抽出条件	別の人が別のツールで再現できる言葉で書く。関数の中に埋まった条件は出典にならない
3 スナップショット	基準日時点の抽出結果。行ごとに、分子・分母・除外のどれに入ったかと、その理由
4 対応の記録	出力した数字ごとに、どの定義書の版とどのスナップショットから出たか、誰が承認したか

注：4つがそろって、出力された数字から元ファイルの行まで一方向に辿れる。出所：Flora株式会社

図表8 出典がない場合に起きること

確かめられない	疑われた数字が正しいのか誤っているのかを判定できない。判定できなければ、訂正の要否も決められない
直せない	誤りの範囲が特定できず、同じ元データから出た他の出力も疑わしくなる
再発防止が書けない	原因が特定できないため、対策が「確認を強化する」という運用の話にしかない
人が代わると失われる	抽出条件が担当者の記憶にある場合、異動で再現できなくなる

注：人的資本の数値は会計監査の対象外で、正確さを担保する仕組みは社内しかない。監査人は「その他の記載内容」として通読し、意見を表明しない。

AIを入れると出力の量が増える。出典の仕組みがないまま出力だけが増えると、確かめられない数字が社内に増える。精度が上がっても、この性質は変わらない。

06 見せてよい範囲

元データの権限設定は、生成された文章には引き継がれない

図表9 生成物で範囲が破れる3つの経路

少人数の集計	部署別・子会社別に切ると、該当者が数名の区分が出る。比率を書いた時点で個人が特定される。女性管理職が2名の部署の「50%」は、誰のことか分かる
読む人の権限	同じテンプレートから作った要約が、自社しか見られない読者にグループ全社の数字を渡す。配信の設定だけでは足りず、生成の時点で範囲を適用する必要がある
目的外の混入	健康・休職・支援制度の利用状況が、登用や配置の説明の中に根拠として現れる。元データに入っていれば、生成は使う

出所：Flora株式会社

図表10 決めておくこと

非表示の閾値	何名未満の集計を表示しないか。画面と出力の両方に同じルールを持たせる
読者ごとの範囲	役割ごとに、どの法人・どの部署・どの粒度まで見てよいか。生成時に適用する
入れない項目	生成の入力に含めない項目を先に決める。健康・休職・支援制度の利用など
記録	誰にどの範囲で何を出したか。後から確認できる形で残す

注：AI事業者ガイドライン（総務省・経済産業省、第1.2版・2026年3月31日）は、人間の判断が形だけになる自動化バイアスに触れ、人間の適切な介在と操作履歴の定期的な確認を求めている。範囲の設計は、その介在が成り立つ前提にあたる。

「後で気をつける」では守れない 非表示や権限は、出力のたびに人が判断すると必ず漏れる。生成の段階でルールとして適用されているかどうか、運用でカバーできる問題とできない問題を分ける。

07 公開事例

公開された失敗事例も、原因は学習データの偏りにあった

図表11 Amazonの採用ツール（2018年に報道、開発は中止）

何をするものか	応募書類を5段階で評価する社内開発の実験的なツール
学習データ	過去10年分の応募書類。技術職の応募者が男性に偏っていた
起きたこと	「women's」を含む記述や、女子大の卒業生を低く評価する傾向が確認された
対応	特定の語について中立化を試みたが、他の形で偏りが再現しうると判断され、開発は中止された
読み取れること	モデルは「過去に採用された人に似た人」を正しく学習した。誤っていたのは、その過去を基準に置いたという前提のほうである

出所：Reuters「Amazon scraps secret AI recruiting tool that showed bias against women」（2018年10月10日）の報道内容による。

図表12 日本の制度が、人に関する判断にAIを使うときに求めていること

根拠	求めていること	人事での意味
AI事業者ガイドライン（第1.2版） 総務省・経済産業省、2026年3月31日	人間の適切な介在。判断が形だけの確認になる自動化バイアスへの注意。操作履歴の定期的な確認	出力をそのまま決定にしないこと、および介在した記録が残ることが前提になる
個人情報保護法 利用目的・開示請求	利用目的をできる限り特定し、本人に通知または公表する。本人からの開示・訂正等の請求に応じる	登用の判断にデータを使うなら、利用目的に含まれている必要がある。開示請求には元データで答える
職業安定法 求職者等の個人情報	収集・保管・使用は、業務の目的の達成に必要な範囲内に限る	採用の場面に直接かかる。在籍者の登用・配置は個人情報保護法の範囲で考える

注：制度の要点を実務の観点から要約したもので、法的助言ではない。出所：総務省・経済産業省「AI事業者ガイドライン（第1.2版）」、個人情報の保護に関する法律、職業安定法
10 © 2026 Flora Co., Ltd.

本資料の計算例との共通点 02・03で示したのは、比較する集団の指定が結果を決めるという話だった。Amazonの事例は、その集団が「過去に採用された人」に固定されていた場合に当たる。**どちらも、モデルは指示どおりに動いている。**

日本企業に当てはめるとき 採用に限らず、登用・配置・育成対象の選定でも同じ構造が起きる。過去の登用実績を学習データにすれば、過去の偏りが再現される。**何を基準に置いたかを書き残し、定期的に見直す**ことが現実的な対処になる。

08 導入前の点検

人事データをAIに渡す前に、12項目

図表13 導入前の点検12項目（本資料の01～07に対応）

- 比較集団が設定として存在する。全社・会社別・等級別・同性内のどれで順位をつけるかを選べる

既定値が何かを知っている。既定のままなら、それを選んだ理由が記録されている

選び方を変えたときの結果を計算した。導入前に、自社の名簿で両方を出して比べた

選から外れた人への説明文がある。「なぜその集団と比べたか」まで答えられる

指標ごとに定義書がある。算出式、範囲、時点、抽出条件、端数処理、改定履歴

出力に定義の版が表示される。数値と定義が一緒に動く

制度改正のときに直す場所が1か所である。定義側の更新が全出力に反映される

元ファイルの行まで遡れる。出力された数字から抽出条件とスナップショットに戻れる

承認の記録がある。誰が、いつ、何を根拠に出したか

少人数の集計が自動で非表示になる。画面と出力の両方で、同じ閾値が効く

読者ごとの範囲が生成の時点で適用される。配信設定ではない。描画時に効く

生成の入力に入れない項目を決めてある。健康・休職・支援制度の利用状況など

注：すべてに印が付かない状態でAIを導入してはいけない、という趣旨ではない。印の付かない項目が、その導入で起こりうる失敗の一覧になる。出所：Flora株式会社

図表14 印が付かなかったときに、最初にやること

付かなかった項目	該当	最初の一步	主担当
比較集団	1～4	自社の名簿で、全社順位と等級別順位の両方を出して人数を比べる。導入判断より前に、選び方の違いを数字で見しておく	人事（要員・登用の担当）
定義	5～7	いま開示している指標の定義書を1枚作る。前任者に抽出条件を聞き取り、届け先ごとに分ける	人事・IR
出典	8・9	次の基準日から、抽出結果のスナップショットを開示値と同じ場所に保存する運用にする	人事・情報システム
見せてよい範囲	10～12	非表示の閾値と、役割ごとに見てよい範囲を決める。生成の入力に入れない項目を先に列挙する	人事・情報システム・法務

注：順番は上から。比較集団と定義が決まっていない段階で出典の仕組みを作っても、何を辿れるようにするかが決まらない。出所：Flora株式会社

METHODOLOGY & SOURCES

本資料の作成方法と出所

作成方法 02と03の計算例は、Flora株式会社が作成した架空のグループ「サンライズHD」（12社・従業員8,420名）のサンプルデータによるものです。実在する企業のデータではありません。等級の中での実力の分布は2社で同じだと置いたうえで、順位をつける集団だけを変えて計算しています。読者が自社の名簿で同じ計算を再現できるよう、前提と数値をすべて記載しました。

なぜ実在企業のデータを使わないか 本資料が示すのは、特定の企業で観測された事実ではありません。算数の性質です。比較集団を変えれば結果が変わることは、どの企業のデータでも同じように起きます。構成された数値で完全に示せる論点であり、実在の従業員のデータを用いる理由がありません。

免責 本資料は、人事データをAIに用いる際の論点を整理したものであり、特定の製品・手法の適法性や有効性を保証するものではありません。個別の導入にあたっては、社内の法務・情報セキュリティ部門にご確認ください。

出所 一次資料

- ・総務省・経済産業省「AI事業者ガイドライン（第1.2版）」（2026年3月31日）
- ・Reuters「Amazon scraps secret AI recruiting tool that showed bias against women」（2018年10月10日）
- ・企業内容等の開示に関する内閣府令 第二号様式 記載上の注意（29）／日本公認会計士協会 監査基準報告書720「その他の記載内容に関連する監査人の責任」
- ・厚生労働省「状況把握、情報公表、認定基準等における解釈事項について」（女性活躍推進法にもとづく指標の算出方法）
- ・Flora株式会社「サンライズHD サンプルデータ」（架空のグループ、2026年9月版）

数値の扱い 本資料に登場する人数・比率はすべてサンプルデータです。公的統計や実在企業の数値と混在させていません。引用される場合は、架空のグループの計算例であることを併記してください。

FLORA / CAPLEAD



AIが言えるのは、データが証明できることと、読む人に見せてよいことだけ

CapLeadは、グループ各社に散在する人事データを、出典まで遡れる一つの基盤で管理するデータ基盤である。開示・分析・AIの3つの出力が、同じ一つのデータと同じ定義を参照する。本資料の01～07で挙げた4つの壊れ方に、製品側で対応している。

①

定義レジストリ 指標ごとに定義を固定して版管理する。制度改正は定義側の更新で全出力に反映される（04に対応）

②

行まで遡れる出典 各社が投入したExcel・CSV・PDFを元ファイルの行まで保持し、出力した数字から戻れる（05に対応）

③

範囲つきの生成 同じテンプレートでも、描画時に読者のスコープを適用する。10名未満の集計は出力にも現れない（06に対応）

分析と開示が同じ数字である 指標エクスプローラ、ドライバー分析、ベンチマーク、経年比較の結果と、有価証券報告書・女性活躍推進法の公表・健康経営度調査に出す数字が、同じデータと同じ定義から出る。経営会議で見た数字と開示した数字が違う状態が起きない。

進め方 30分ヒアリング（無償）→ NDA後、前年度の実データで検証 → 予算サウンディング → 定義すり合わせ・初期取込 → 本稼働。費用は対象法人数と開示先により個別のお見積もり。画面例はサンプルデータ・架空のグループによるものです。

CONTACT

まず30分、貴社の人事データで何ができるかをお聞かせください。

いま持っているデータの状態を伺い、AIに渡せる状態までに何が要るかをその場でお伝えします。

SERVICE

flora-tech.jp

CONTACT

admin@floramaternity.com

執筆

Flora株式会社 CapLeadチーム 2026年9月

本資料の内容の無断転載を禁じます。引用の際は「Flora株式会社『人事AIの失敗パターン 2026』」と出所を明記してください。本資料の計算例は架空のグループのサンプルデータによるもので、実在の企業・団体・人物とは関係ありません。2026年9月時点の情報にもとづいて作成しています。