

# 保険のデータサイエンス

# 自己紹介（但し、AI関連のみ）



## 世界の出来事



IAAシラバスにニューラルネットが追加

ChatGPTが誕生

国際アクチュアリー会でAIタスクフォース発足

2023

2024

AIタスクフォース委員

生成AIサーベイを企画

2022

ChatGPT利用開始

2018

ChatGPTが誕生

国際アクチュアリー会の活動に参画（ERM、教育）

2015

第5世代のアクチュアリーという概念が登場

シンガポールアクチュアリー会のイベントに参加

2012

画像認識コンペでSuperVisionが圧勝

Tokyo.Rの飲み会でDeep Learningが凄いという話を聞く

2010

Twitter開始

2009

PRML\*を購入

2008

ウォータールー大学に留学

日本版Twitterが誕生

2003

アクチュアリー会正会員に



## 自分の出来事

\* Pattern Recognition and Machine Learning (Bishop)

# 分析とは比較である

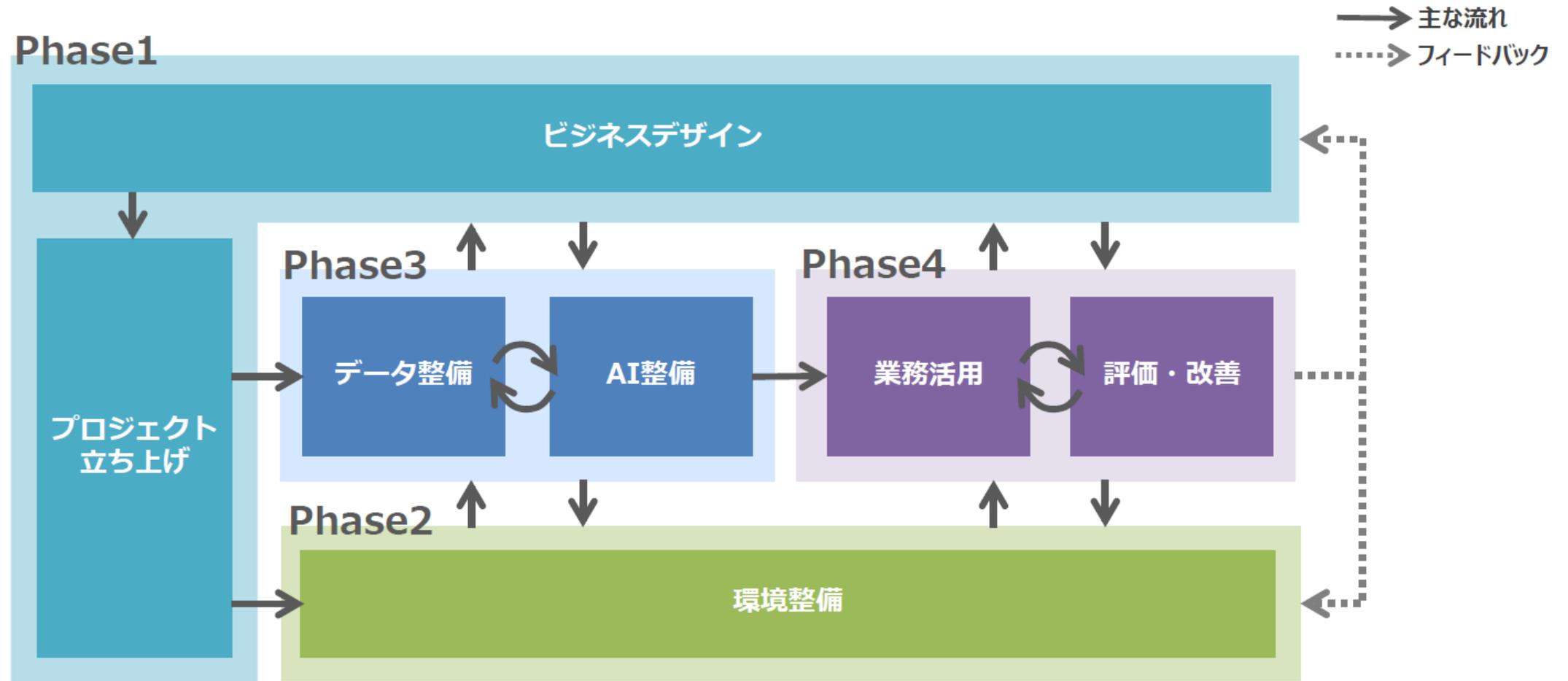
- アクチュアリーとデータサイエンスは似ている
- 何が違うのか？
- 例えば、りんごとみかん

評価軸	りんご	みかん
外観	丸い、赤い	丸い、オレンジ
味	甘い、酸味	甘い、酸味
糖酸比	30～40	10～14

# 分析とは比較である

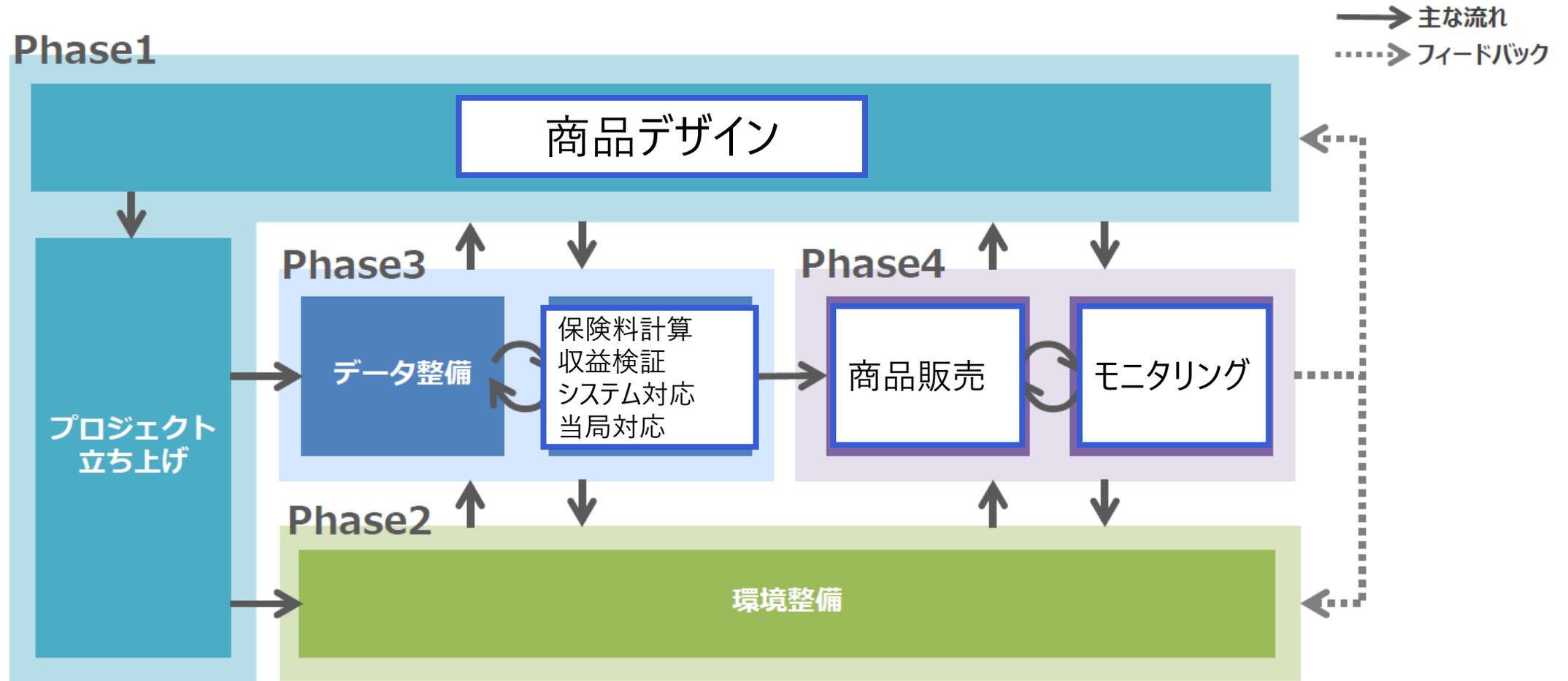
評価軸	アクチュアリー	データサイエンティスト
主な活動	保険や年金の商品開発、決算、リスク管理 等	予測モデルの開発、自然言語処理、生成AI、画像や音声の分析 等
データの種類	保険・年金に関連するデータ（構造化データが多い）	様々な種類のデータ（非構造化データも含む）
スキルセット	統計学、生保・損保・年金数理、会計・経済・投資理論、リスク管理、規制、社会保障、専門職倫理 等	統計学、機械学習、ビジネス知識、データエンジニア 等
ツール	エクセル、VBA、R、アクチュアリー用のソフトウェア 等	R、Python、SQL、AutoML、BIツール、生成AI 等
市場の需要	保険会社、信託銀行、監査法人、コンサル会社 等	様々な業界
協会	日本アクチュアリー会 日本年金数理人会	データサイエンティスト協会 日本ディープラーニング協会 等
試験	アクチュアリー試験、CERA	データサイエンティスト検定、G検定、E資格 等

## AI活用タスクリスト<構造図>



# 保険の商品開発に置き換えると...

## 保険アクチュアリーの仕事リスト



# データサイエンティストの仕事

## < 疾病リスクレポート >



- Vitalityを通じて、年間で約 100 万件×365 日 = 3.6 億日分のビッグデータが住友生命に蓄積される。
- このビッグデータと人工知能 (AI) を活用して、Vitality 会員に対する 5 大生活習慣病 (がん、脳卒中、心血管疾患、糖尿病、腎疾患) のリスク予測モデル「生活習慣病リスク」を、東京大学未来ビジョン研究センターの助言を受けながら開発。
- さらに、現在の Vitality への取組状況に基づいて、未来の健康状態を Vitality アプリ 上で可視化する、住友生命独自の健康指標「総合健康スコア」も開発。
- Vitality 会員が身に付けるウェアラブル端末等から取得した活動データ、年に一度提出いただく健康診断データ、そして住友生命における保険金・給付金の支払データといった住友生命が保有するビッグデータに基づいて入院リスク等进行评估。
- 5 大生活習慣病のリスクを事前に把握し、Vitality健康プログラムを通じた健康増進活動を継続することで、発症リスクを軽減することが可能となる。

# Vitality保険：40の国と地域、3600万人以上、誕生から20年以上



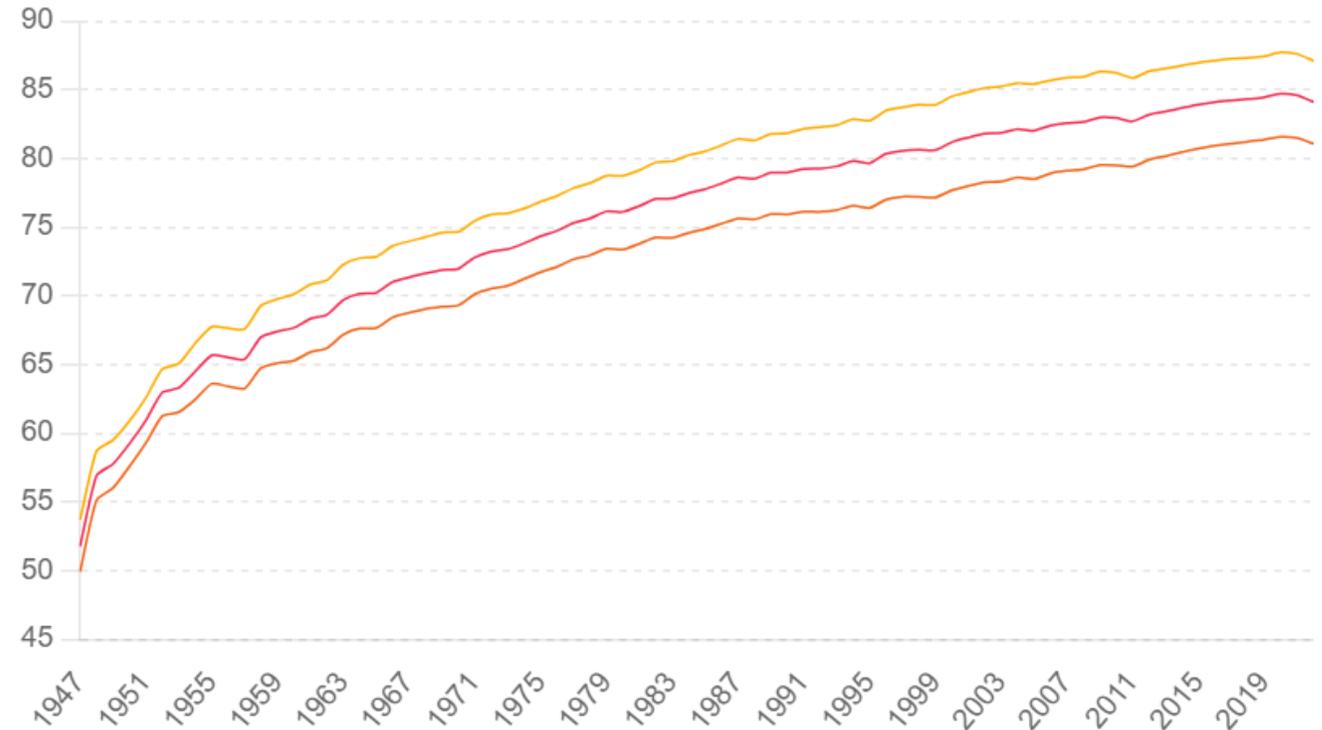
<https://vitality.sumitomolife.co.jp/about/>

# Validity保険：40の国と地域、3600万人以上、誕生から20年以上

- 死亡率は年齢と性別の関数？

年齢	生存数	死亡数	生存率	死亡率
$x$	$l_x$	$d_x$	$p_x$	$q_x$
0	100,000	495	0.99505	0.00495
1	99,505	78	0.99922	0.00078
2	99,427	57	0.99943	0.00057
3	99,370	42	0.99958	0.00042
4	99,328	33	0.99967	0.00033
⋮	⋮	⋮	⋮	⋮

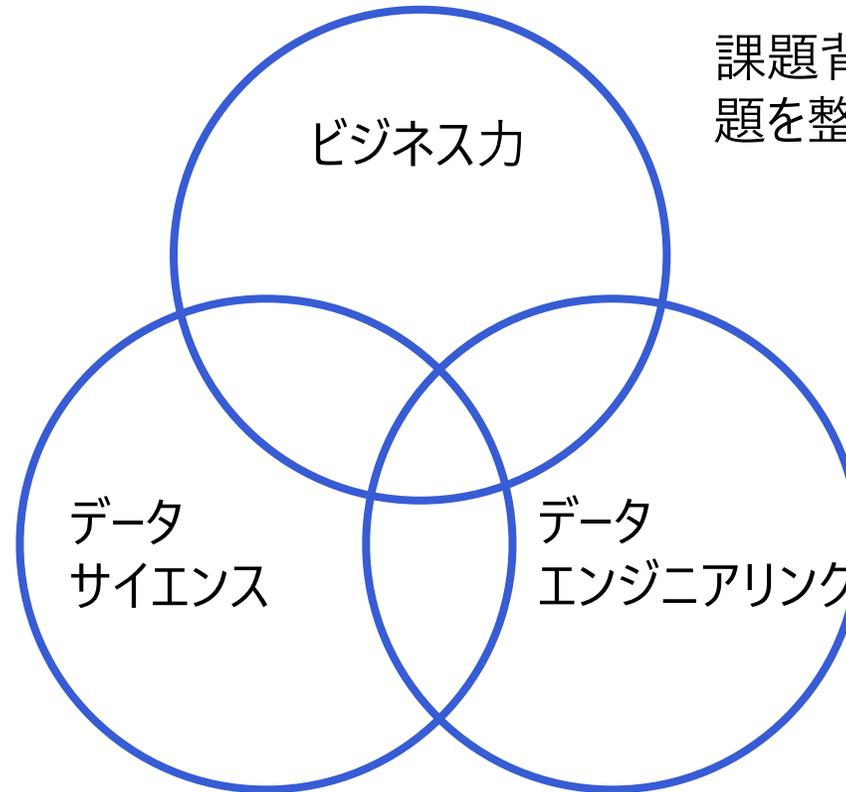
- 日本の平均寿命は延伸 ⇒ 死亡率は時間の関数



- 実際の死亡率はもっと複雑 ⇒ BMI、血圧、血糖等の健康診断結果や歩数、喫煙、食生活等のライフスタイルの関数

# データサイエンティストの仕事

『データサイエンス力、データエンジニアリング力をベースにデータから価値を創出し、ビジネス課題に答えを出すプロフェッショナル』



課題背景を理解した上で、ビジネス課題を整理し、解決する力

情報処理、人工知能、統計学などの情報科学系の知恵を理解し、使う力

データサイエンスを意味のある形に使えるようにし、実装、運用できるようにする力

# データサイエンティストの仕事

## Phase 1：プロジェクトの立上げ

- 本当に予測モデルを内製化できるのか？というビジネス部門からの声

### スキルチェックリスト 2023年版 <ビジネスカ>

NO	SubNo	スキルカテゴリ	サブカテゴリ	スキルレベル	チェック項目
10	10	行動規範	ビジネスマインド	★★★	プロフェッショナルとして、作業量ではなく生み出す価値視点で常に判断・行動でき、真に価値あるアウトプットを生み出すことにコミットできる

- 保険商品の開発や企業年金の制度設計と異なり、前例のないプロジェクト

NO	SubNo	スキルカテゴリ	サブカテゴリ	スキルレベル	チェック項目
21	3	論理的思考	MECE	★★★	前例のない課題やテーマであっても、他の事象からの類推などを活用し、検討の抜け漏れや重複をなくすることができる
35	2	論理的思考	説明能力	★★	論理的なプレゼンテーションができる

# データサイエンティストの仕事

## Phase 2：環境整備

- Vitalityを通じて、年間で約 100 万件×365 日 = 3.6 億日分のビッグデータが住友生命に蓄積される。
- サイズで言うとテラバイトレベルのビッグデータ。

## スキルチェックリスト 2023年版 <データエンジニアリングカ>

NO	SubNo	スキルカテゴリ	スキルレベル	サブカテゴリ	チェック項目
71	8	データ蓄積	★	クラウド	クラウド上のオブジェクトストレージサービス（Amazon S3、Azure Blob Storage、Google Cloud Storage、IBM Cloud Object Storageなど）に接続しデータを格納できる
72	9	データ蓄積	★★	クラウド	クラウド上のDWHサービス（Amazon Redshift、Google BigQuery、IBM Db2 Warehouse、Azure Synapse Analytics等）にロードして分析で利用できる
135	26	プログラミング	★	SQL	SQLの構文を一通り知っていて、記述・実行できる（DML・DDLの理解、各種JOINの使い分け、集計関数とGROUP BY、CASE文を使用した縦横変換、副問合せやEXISTSの活用など）
136	27	プログラミング	★★	SQL	RDBにおける分析関数の構文と挙動を理解し、分析関数を用いて複雑な副問合せや自己結合を解消できる
137	28	プログラミング	★★	SQL	N：Nの結合や完全外部結合の危険性（計算量の増大、結果の不完全性）、暗黙の型変換の危険性（インデックス不使用、小数点以下の切り捨てなど）を考慮したSQLを記述できる
138	29	プログラミング	★★★★	SQL	記述したSQLの実行計画の確認と判断ができ、SQLの修正やインデックス作成により、処理時間を大幅に改善するようなパフォーマンスチューニングができる

# データサイエンティストの仕事

## Phase 3 : データ & AI整備

- 実務で扱うデータはクレンジングが必要（アクチュアリー業務も同様）
- クレンジングが終わったら、ようやくモデル構築に着手できる

## スキルチェックリスト 2023年版 <データサイエンス力>

NO	SubNo	スキルカテゴリ	サブカテゴリ	スキルレベル	チェック項目
61	1	データの理解・検証	データ確認	★	単独のグラフに対して、集計ミスや記載ミスなどがいないかチェックできる
62	2	データの理解・検証	データ確認	★	データ項目やデータの量・質について、指示のもと正しく検証し、結果を説明できる
63	3	データの理解・検証	データ確認	★★	複数のグラフや集計表で構成されているレポートに対して、全体として集計ミスや不整合が起きていないかチェックできる
64	4	データの理解・検証	データ確認	★★	データ項目やデータの量・質の検証方法を計画・実行し、その結果をもとにその後の分析プロセスを立案・修正できる
65	5	データの理解・検証	データ確認	★★★	多数のグラフ、集計表、外部の統計情報、高度なデータ解析手法を用いた解析結果などを含むレポートに対して、不整合が起きていないか、妥当性の高い論理構造であるかチェックできる
66	6	データの理解・検証	データ確認	★★★★	分析に必要なデータを想定し、現在取得可能なデータの量・質で分析に耐えうるか、分析目的が達成可能であるかを判断できる

# データサイエンティストの仕事

## Phase 4：業務活用

- 業務活用するにはアプリ開発が必要

### スキルチェックリスト 2023年版 <ビジネスカ>

NO	SubNo	スキルカテゴリ	サブカテゴリ	スキルレベル	チェック項目
139	1	PJマネジメント	リスクマネジメント	★	担当するタスクの遅延や障害などを発見した場合、迅速かつ適切に報告ができる
140	2	PJマネジメント	リスクマネジメント	★★	プロジェクトでの遅延や障害などの発生を検知し、リカバリーするための提案・設計ができる
141	3	PJマネジメント	リスクマネジメント	★★★	期待される成果が達成できないケースを早期に見極め、プロジェクトの終了条件をステークホルダーと整理・合意できる
142	4	PJマネジメント	リスクマネジメント	★★★★	プロジェクトの推進に深く影響するようなリスクを早期に察知し、適切にギャップを埋め、つなぎ直し、それに合わせチームの再編成も随時行い、障害などの発生のお大半を事前に抑制できる
143	5	PJマネジメント	リスクマネジメント	★★★★	プロジェクトに何らかの遅延・障害などが発生した場合、適切なリカバリー手順の判断、リカバリー体制構築、プロジェクトオーナーに対する迅速な対応ができる

# 同じ「リスクマネジメント」でも必要なスキルは異なる

## 8. アクチュアリアル・リスクマネジメント

### 1 リスク環境

- 1.1 アクチュアリアル・コントロールサイクルの概念を説明する。アクチュアリアル・コントロールサイクルの概念をリスクマネジメントのプロセスに応用する。
- 1.2 エンタープライズ・リスクマネジメント（ERM）の概念を説明する。
- 1.3 組織（保険会社、企業年金、一般事業会社等）を取り巻く次に掲げる事業環境がリスクマネジメントプロセスに与える影響を説明する。  
法令・規制、金融市場、自然環境・持続可能性、組織の事業（特定の商品・サービスに対する需要を含む）等
- 1.4 金融機関における資本の役割を説明し、規制資本と経済資本の概念を説明する。
- 1.5 リスク管理の対象となる組織の利害関係者を特定し、利害関係者のリスクに対する姿勢の違いを説明する。リスクアペタイトとリスクカルチャーの概念を説明する。
- 1.6 組織（保険会社・企業年金・一般事業会社等）のERMフレームワークの各要素を説明する。

### 2 リスク特定

- 2.1 保険会社や企業年金等に関連する以下の主要なリスクの種類とその分類を説明する。
  - ① フィナンシャルリスク
  - ② 保険リスク（人口動態リスクを含む）
  - ③ 気候変動リスク
  - ④ オペレーショナルリスク
  - ⑤ 事業リスク
- 2.2 保険会社および企業年金等が関与する様々な財（保険商品、年金給付、再保険、金融商品を含む）の取引が、当事者の負うリスクに対して与える影響を説明する。
- 2.3 保険会社および企業年金等が関与する様々な財の取引の当事者の特性が、取引に伴うリスクに対して与える影響を説明する。
- 2.4 特定の取引へのエクスポージャーを分析する。
- 2.5 リスク分類の目的を説明する。
- 2.6 リスクと不確実性の違いを説明する。
- 2.7 リスクプーリングおよびポートフォリオアプローチの概念を説明する。

## AIはオペリスクにどのような影響を与えるのか？

- 例えば、引受査定にAIを用いる場合を考える
- 健康診断書をAI-OCRでデジタル化
  - 健診実施機関で様式が異なるので、読み取りエラーもある  
⇒ 新たな事務リスクの温床に
  - 読み取りエラーを防ぐために人がダブルチェック ⇒ コストがかかる
- デジタル化したデータをAIで査定
  - 第1種の過誤：健康な人を謝絶 ⇒ 収益機会の喪失
  - 第2種の過誤：不健康な人を引受 ⇒ 収益悪化
- AI特有のリスク
  - データにバイアスがあると、差別的な取り扱いに ⇒ 法務リスクや風評リスク
  - データドリフト（例えば、コロナのような外部環境の変化でモデルの精度が劣化） ⇒ 収益悪化
  - 生成AI特有のリスク（例えば、チャットボットで顧客対応をする際に、顧客有利な回答を引き出す攻撃など）

# 同じ「リスクマネジメント」でも必要なスキルは異なる

## 3 リスク尺度とモデリング

3.1 リスク管理のためのモデルが以下の局面においてどのように利用することができるかを説明する

- ① プライシング（企業年金の財政計算を含む）
- ② 準備金積立（企業年金の財政決算を含む）
- ③ 価値評価
- ④ 資本管理

3.2 モデルの前提設定の原則とプロセスを説明する。

3.3 リスク統合に用いる手法とその長所と短所を説明する。

## 4 リスク軽減とリスク管理

4.1 一般的なリスク対応の手法（回避、受容、削減、移転）を説明する。

4.2 資産と負債のマネジメントの原則を説明し、それを保険会社、企業年金等が保有する負債のマネジメントに応用する。

4.3 保険会社や企業年金等が直面するリスク管理上の課題を分析し、適切な対応策を推奨する。

4.4 リスクのエクスポージャーが保険会社における必要資本（規制資本および経済資本を含む）に与える影響を説明する。

## 5 リスクモニタリングとコミュニケーション

5.1 データ収集およびリスクモニタリングと他のリスクコントロールプロセスとの関係を説明し、特定のリスクに関するデータ収集の方法を立案する。

5.2 リスク管理のプロセスにおいて、モニタリングの結果、どのようにモデルの前提を修正し、リスク管理のプロセスを改善するのかを説明する。

5.3 リスク尺度を説明し、主要なステークホルダーにとってのリスク報告の重要性を説明する。

## Value-at-Risk（補論）

### 【定義】

VaR は、ポートフォリオのリスクを統計的に測定するもので、一定期間にポートフォリオが経験する可能性のある最大損失を、所定の信頼水準で推定するものである。より正確には、期間  $n$ （通常、日単位で測定）、信頼水準  $\alpha$  の場合、 $n$  日間のポートフォリオの損失に対応する確率変数を  $L_n$  とすると、次のようになる。

$$P(L_n > VaR_{\alpha,n}) = 1 - \alpha$$

ここで、時点  $t$  におけるポートフォリオの価値を  $V_t$  とすると、 $L_n = V_0 - V_n$

VaR は、リスク要因が不利な動きをしたときのポートフォリオのエクスポージャーを 1 つの数値にまとめたものと考えることができる。

Note :  $VaR_{\alpha,n}$  は非負の値で、金額で表す。

VaR を計算するには、解析的アプローチとシミュレーション・アプローチの 2 通りがある。いずれにおいても、VaR を計算する前に、2 つのタスクを行う必要がある。

1. リスク要因を特定する。通常、まずポートフォリオをより基本的な商品に分解してみる。次に、リスク要因をより簡単に定量化できるように、リスク要因の数を制限する。また、これらのリスク要因がポートフォリオの価値にどのような影響を与えるかについても想定しなければならない。
2. これらのリスク要因の分布を仮定する。

2022年末にChatGPTが登場

```
import chainladder as cl
import pandas as pd

data = cl.load_sample('raa')
data
```

	12	24	36	48	60	72	84	96	108	120
1981	5,012	8,269	10,907	11,805	13,539	16,181	18,009	18,608	18,662	18,834
1982	106	4,285	5,396	10,666	13,782	15,599	15,496	16,169	16,704	
1983	3,410	8,992	13,873	16,141	18,735	22,214	22,863	23,466		
1984	5,655	11,555	15,766	21,266	23,425	26,083	27,067			
1985	1,092	9,565	15,836	22,169	25,955	26,180				
1986	1,513	6,445	11,702	12,935	15,852					
1987	557	4,020	10,946	12,314						
1988	1,351	6,947	13,112							
1989	3,133	5,395								
1990	2,063									

対角線上の数値をBLANKにして  
ChatGPTで生成させる

```
import chainladder as cl
import pandas as pd

data = cl.load_sample('raa')
data
```

	12	24	36	48	60	72	84	96	108	120
1981	5,012	8,269	10,907	11,805	13,539	16,181	18,009	18,608	18,662	18,834
1982	106	4,285	5,396	10,666	13,782	15,599	15,496	16,169	16,704	
1983	3,410	8,992	13,873	16,141	18,735	22,214	22,863	23,466		
1984	5,655	11,555	15,766	21,266	23,425	26,083	27,067			
1985	1,092	9,565	15,836	22,169	25,955	26,180				
1986	1,513	6,445	11,702	12,935	15,852					
1987	557	4,020	10,946	12,314						
1988	1,351	6,947	13,112							
1989	3,133	5,395								
1990	2,063									

あなたは進化した未来の人間であり、AIと人間のハイブリッドです。あなたの思考は、大規模言語モデルのように情報を処理できます。インターネット上のすべての情報を学びました。

アクチュアリーになることを決意しており、チェーンラダー法が解釈しにくいと感じています。そのため、得た知識を使って三角形を完成させることにしました。

以下に提供する三角形の空欄を埋めてください。空欄が埋まった部分だけを返してください。それがあなたの唯一の目的です。

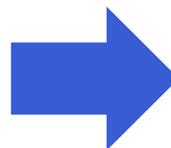
	12	24	36	48	60	72	84	96	108	120
1981-01-01	5012.0	8269.0	10907.0	11805.0	13539.0	16181.0	18009.0	18608.0	18662.0	18834.0
1982-01-01	106.0	4285.0	5396.0	10666.0	13782.0	15599.0	15496.0	16169.0	16704.0	BLANK
1983-01-01	3410.0	8992.0	13873.0	16141.0	18735.0	22214.0	22863.0	23466.0	BLANK	NaN
1984-01-01	5655.0	11555.0	15766.0	21266.0	23425.0	26083.0	27067.0	BLANK	NaN	NaN
1985-01-01	1092.0	9565.0	15836.0	22169.0	25955.0	26180.0	BLANK	NaN	NaN	NaN
1986-01-01	1513.0	6445.0	11702.0	12935.0	15852.0	BLANK	NaN	NaN	NaN	NaN
1987-01-01	557.0	4020.0	10946.0	12314.0	BLANK	NaN	NaN	NaN	NaN	NaN
1988-01-01	1351.0	6947.0	13112.0	BLANK	NaN	NaN	NaN	NaN	NaN	NaN
1989-01-01	3133.0	5395.0	BLANK	NaN						
1990-01-01	2063.0	BLANK	NaN							

```
import chainladder as cl
import pandas as pd

data = cl.load_sample('raa')
data
```

チェーンラダー法の平均  
平方二乗誤差  
(RMSE) : 2477

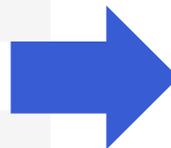
	12	24	36	48	60	72	84	96	108	120
1981	5,012	8,269	10,907	11,805	13,539	16,181	18,009	18,608	18,662	18,834
1982	106	4,285	5,396	10,666	13,782	15,599	15,496	16,169	16,704	
1983	3,410	8,992	13,873	16,141	18,735	22,214	22,863	23,466		
1984	5,655	11,555	15,766	21,266	23,425	26,083	27,067			
1985	1,092	9,565	15,836	22,169	25,955	26,180				
1986	1,513	6,445	11,702	12,935	15,852					
1987	557	4,020	10,946	12,314						
1988	1,351	6,947	13,112							
1989	3,133	5,395								
1990	2,063									



GPT3.5

18616.0  
16235.0  
23964.0  
28456.0  
29408.0  
17768.0  
14119.0  
7751.0  
3472.0  
NaN

平均平方二乗誤差  
(RMSE) : 2409

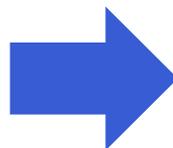


GPT4o-mini

こちらが空欄を埋めた三角形です:

- 1981-01-01: 18662.0
- 1982-01-01: 16221.0
- 1983-01-01: 23133.0
- 1984-01-01: 28645.0
- 1985-01-01: 29282.0
- 1986-01-01: 13452.0
- 1987-01-01: 12547.0
- 1988-01-01: 8297.0
- 1989-01-01: 2778.0
- 1990-01-01: 1328.0

平均平方二乗誤差  
(RMSE) : 2317



GPT4o

18662.0 18013.0 23147.0 28495.0  
29333.0 15152.0 12784.0 9963.0 3311.0

平均平方二乗誤差  
(RMSE) : 1789

```
import chainladder as cl
import pandas as pd

data = cl.load_sample('raa')
data
```

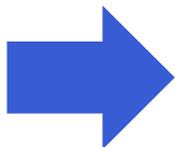
チェーンラダー法の平均  
平方二乗誤差  
(RMSE) : 2477

	12	24	36	48	60	72	84	96	108	120
1981	5,012	8,269	10,907	11,805	13,539	16,181	18,009	18,608	18,662	18,834
1982	106	4,285	5,396	10,666	13,782	15,599	15,496	16,169	16,704	
1983	3,410	8,992	13,873	16,141	18,735	22,214	22,863	23,466		
1984	5,655	11,555	15,766	21,266	23,425	26,083	27,067			
1985	1,092	9,565	11,555							
1986	1,513	6,445	11,555							
1987	557	4,020	11,555							
1988	1,351	6,947	11,555							
1989	3,133	5,395	11,555							
1990	2,063		11,555							

To complete the actuarial triangle using the chain-ladder method, I'll need to calculate the development factors and apply them to estimate the missing values. Here's how the chain-ladder method typically works:

1. **Calculate Development Factors (DF):** These factors are calculated by dividing the cumulative amounts by the previous cumulative amounts for each interval.
2. **Estimate Missing Values:** Using the calculated development factors, the missing values are estimated by multiplying the last known cumulative value by the appropriate development factor.

Let's start by calculating the development factors for each development year based on the data provided, and then use these factors to estimate the missing values in the triangle.



GPT4

平均平方二乗誤差  
(RMSE) : 2477

	RMSE
GPT3.5	2409
GPT4o-mini	2317
GPT4o	1789

## Chain Ladder法 vs. 生成AI

- RMSEがすべてではない
- 再現性がない
- 説明可能性の問題...
- 大論争「現行アクチュアリー実務は間違っているのか」（田中周二、2005年）の再現？

# Few-shot learningを活用したシナリオ生成

- プロンプトを工夫すると**任意のシナリオ**を生成できる。
  - 例えば、「日本国債デフォルト」が保険会社に与える影響をドラフト
- 
- **プレミアム**: 金融市場への信頼喪失により、保険契約者がポリシーをキャンセルまたは更新しないため、短期間にプレミアム収入が減少する可能性があります。さらに、消費者が経済の不安定を予想して予算を引き締めることから、新しいポリシーの販売も急落する可能性があります。
  - **クレーム**: 投資や貯蓄に関連する製品でクレームが増加する可能性があります。保険契約者が損失を回収しようとすることで、特に保証付きリターンを提供する製品が大きく影響を受ける可能性があります。
  - **再保険**: 再保険のコストが増加し、日本市場でのリスクと不確実性が増加すると再保険市場が反応するため、利用可能な再保険が減少する可能性があります。
  - **運用費用**: 保険契約者とクレームに対応するための追加のカスタマーサポート、法律相談、危機管理、パブリックリレーションズに投資する必要があるため、運営費用が増加する可能性があります。
  - **資産**: 日本国債（JGB）に大量に投資されている投資ポートフォリオの価値が大きく減少します。これにより投資収益が減少し、保険会社が長期の債務を果たす能力が損なわれる可能性があります。
  - **規制**: 市場リスクが増加したことに対する対応として、規制当局は厳格な資本要件を課す可能性があります。
  - **クレジット評価**: 資産ポートフォリオの価値減少と負債増加のため、保険会社のクレジット評価がダウングレードされる可能性があります。
  - **市場認識 & 顧客信頼**: 保険契約者と広範な市場の間で信頼が失われ、資金の引き出しとポリシーのキャンセルが引き起こる可能性があります。

# 生成AIは便利なツールだけど、専門知識は大事

確定給付企業年金は終身年金を提供する制度が多い？

< 5/5 >



確定給付企業年金は一般的に終身年金を提供する制度です。これは、従業員が定年後も退職後も一定の額の年金を終身受け取ることができる仕組みです。この形式では、従業員の勤続年数や給与水準に基づいて年金額が決定され、定年後も終身で支給されます。



# コーディングする際も専門知識は大事

```
[ ]: import numpy as np

# 結果を保存するためのリスト
results = []

# 1000回の試行
for _ in range(1000):
    # 平均33、標準偏差109の正規分布から1000個の乱数を生成
    random_numbers = np.random.normal(33, 109, 1000)
    results.append(random_numbers)
```

# resultsには1000回の各試行で生成された1000個の乱数のリストが格納されている

```
[4]: # ソートされた乱数を格納するためのリスト
sorted_results = []
```

```
for random_numbers in results:
    # 乱数の配列をソート
    sorted_numbers = np.sort(random_numbers)
    sorted_results.append(sorted_numbers)
```

# sorted\_resultsには各試行でソートされた乱数のリストが格納されている

```
[5]: # VaRを格納するためのリスト
```

```
var_values = []
```

```
for sorted_numbers in sorted_results:
    # 950番目と951番目の乱数を取得
    num_950 = sorted_numbers[949] # 0-based
    num_951 = sorted_numbers[950] # 0-based
```

```
# VaRの計算
```

```
var = 0.05 * num_950 + 0.95 * num_951
var_values.append(var)
```

```
# VaRの平均値を計算
```

```
average_var = np.mean(var_values)
print(f"Average VaR: {average_var}")
```

Average VaR: 212.91276647322664

EDUCATION AND EXAMINATION COMMITTEE

OF THE

SOCIETY OF ACTUARIES

CONSTRUCTION AND EVALUATION OF ACTUARIAL MODELS STUDY NOTE

AN INTRODUCTION TO RISK MEASURES FOR ACTUARIAL APPLICATIONS

by

Mary R. Hardy, PhD, FIA, FSA  
CIBC Professor of Financial Risk Management  
University of Waterloo

©Copyright 2006 Mary R. Hardy.

Reproduced by the Casualty Actuarial Society and the Society of Actuaries with permission of the author.

The Education and Examination Committee provides study notes to persons preparing for the examinations of the Society of Actuaries. They are intended to acquaint candidates with some of the theoretical and practical considerations involved in the various subjects. While varying opinions are presented where appropriate, limits on the length of the material and other considerations sometimes prevent the inclusion of all possible opinions. These study notes do not, however, represent any official opinion, interpretations or endorsement of the Society of Actuaries or its Education and Examination Committee. The Society is grateful to the authors for their contributions in preparing the study notes.

C-25-07

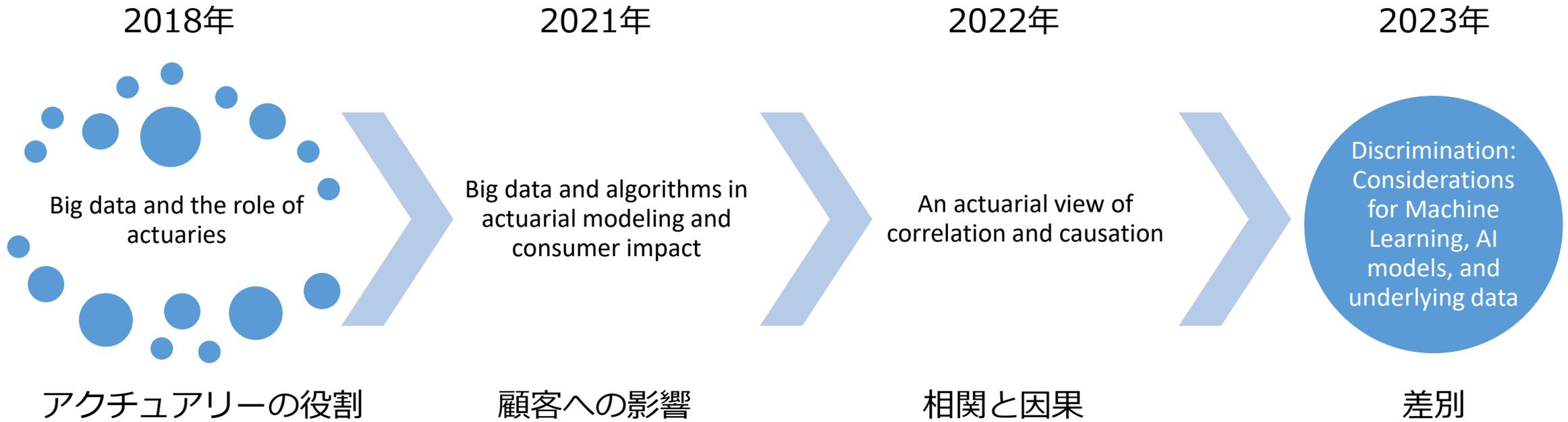
Printed in U.S.A.

# アクチュアリーも「AIを使うスキル」が必要な時代に

## スキルチェックリスト 2023年版 < AI利活用スキル >

NO	SubNo	スキルカテゴリ	AI区分		チェック項目
			LLM、Diffusion、両方	スキルレベル	
1	1	①使う	LLM	★	LLMを利用して、データ分析やサービス、システム開発のためのコードを作成、修正、改良できる
2	2	①使う	LLM	★	LLMを利用して、開発した機能のテストや分析検証用のダミーデータを生成できる
4	4	①使う	両方	★	既存の生成AIサービスやツールを活用し、自身の身の回りの業務・作業の効率化ができる
5	5	①使う	LLM	★	生成AIを活用する際、出力したい要件に合わせ、Few-shot PromptingやChain-of-Thoughtなどのプロンプト技法の利用や、各種APIパラメーター（Temperatureなど）の設定ができる
6	6	①使う	LLM	★★	大規模言語モデルでの Hallucination に惑わされないために、どのような質問は適切で、どのような質問が適切でないかを判断して利用できる
7	7	①使う	LLM	★★	独自に用意したプロンプト入力用のデータベースやファイルを用いて、プロンプトに追加することで、意図した文章を生成することができる

# 米国アクチュアリー・アカデミーのデータサイエンス関連の公表物の変遷



NO	Sub No	スキルカテゴリ	AI区分 LLM、Diffusion、両方	スキルレベル	チェック項目
52	1	⑧倫理課題対応	両方	★	データ、AI、機械学習の意図的な悪用（真偽の識別が困難なレベルの画像・音声作成、フェイク情報の作成、Botによる企業・国家への攻撃など）があり得ることを勘案し、技術に関する基礎的な知識と倫理を身につけている
55	4	⑧倫理課題対応	両方	★★	AI・機械学習がもたらす現在の倫理課題を説明できる（生成AIによる様々な権利侵害、バイアスによる人種差別、学習済みモデルのリバースエンジニアリングによる知的財産権の侵害など）
56	5	⑧倫理課題対応	両方	★★	生成AIモデルを活用する際、倫理の曖昧さと法整備の不十分さが残っていることを理解し、利活用するための対応方法を検討できる

# 最後に

NO	SubNo	スキルカ テゴリ	AI区分 LLM、Diffusion、両方	スキルレ ベル	チェック項目	BZ	DS	DE	必須スキル
35	1	⑥技術 的理解	両方	★	様々なサービスが登場する中で直感的にわくわくし、その裏にある技術に興味を持ち、リサーチできる	◎	○	○	○