



ML Enablement Workshop: ML Input module #3

# ML活用組織への シフト

久保 隆宏

Developer Relation  
Machine Learning

# 本コースのアウトライン

## MLの入門

MLの可能性

AI/MLとは

MLの活用事例

MLの技術的制約

## MLプロジェクトの計画

ユースケースの特定

ML拡大の要件

データの要件

期間・実用化の要件

## ML活用組織へのシフト

堅牢なAI戦略

データ戦略

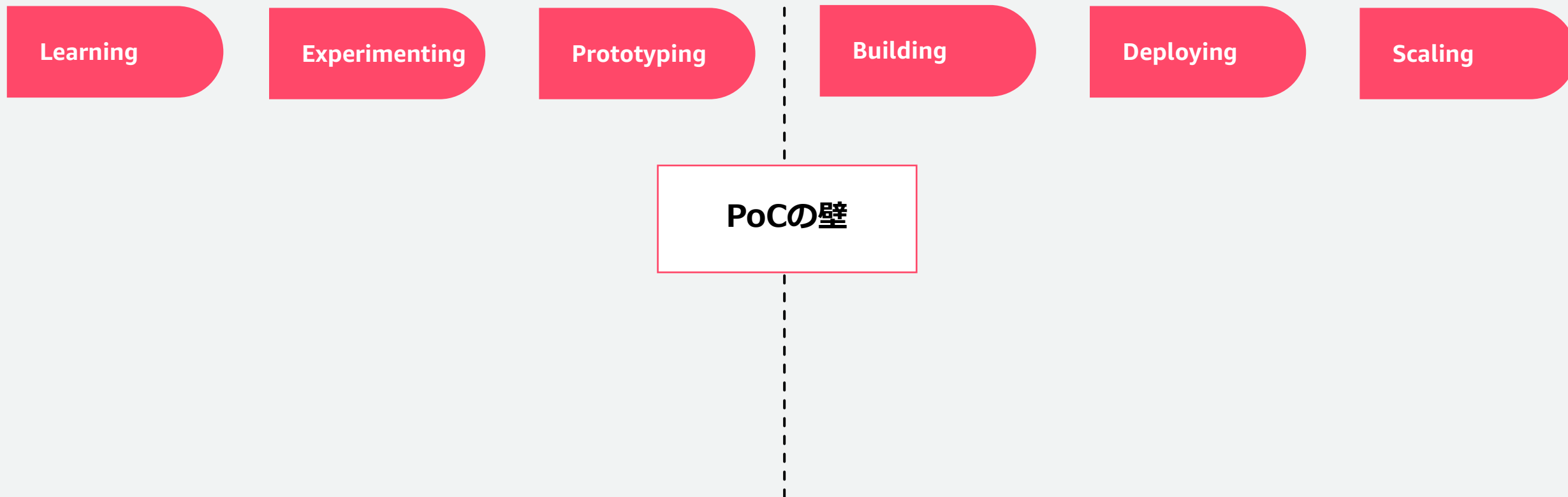
組織内のコラボレーション

推進のためのリーダーシップ

# 組織におけるMLジャーニー

機械学習を学び、検証し、試作する

機械学習モデルを構築・運用・スケールする



# MLジャーニーの推進に必要な要件

AI活用へのシフトを成功させるために必要な組織としての要件



堅牢なAI戦略



データ戦略



組織内のコラボレーション

これらを推進するためのリーダーシップ

# 1. 堅牢なAI戦略

## 2. データ戦略

## 3. 組織内のコラボレーション

## 4. 推進のためのリーダーシップ

# AI戦略における主な検討事項



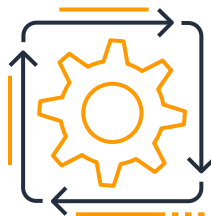
## 適切な問題の 特定

MLによって解決できる適切な問題を定め、組織内のモメンタムを作る。



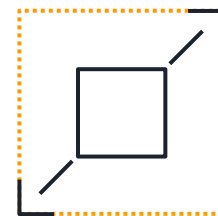
## クラウドの力を 活用

クラウドの活用により、PoCの開発を加速させる。



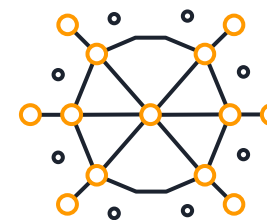
## 前向きな失敗の 奨励

MLプロジェクトに失敗はつきもの。失敗を次の機会のための学びとして扱う。



## PoCを越えるス ケーラビリティ

PoCの先を見据えたイメージを作る。新たなプロジェクトを高スケールで実用化する計画を立てる。

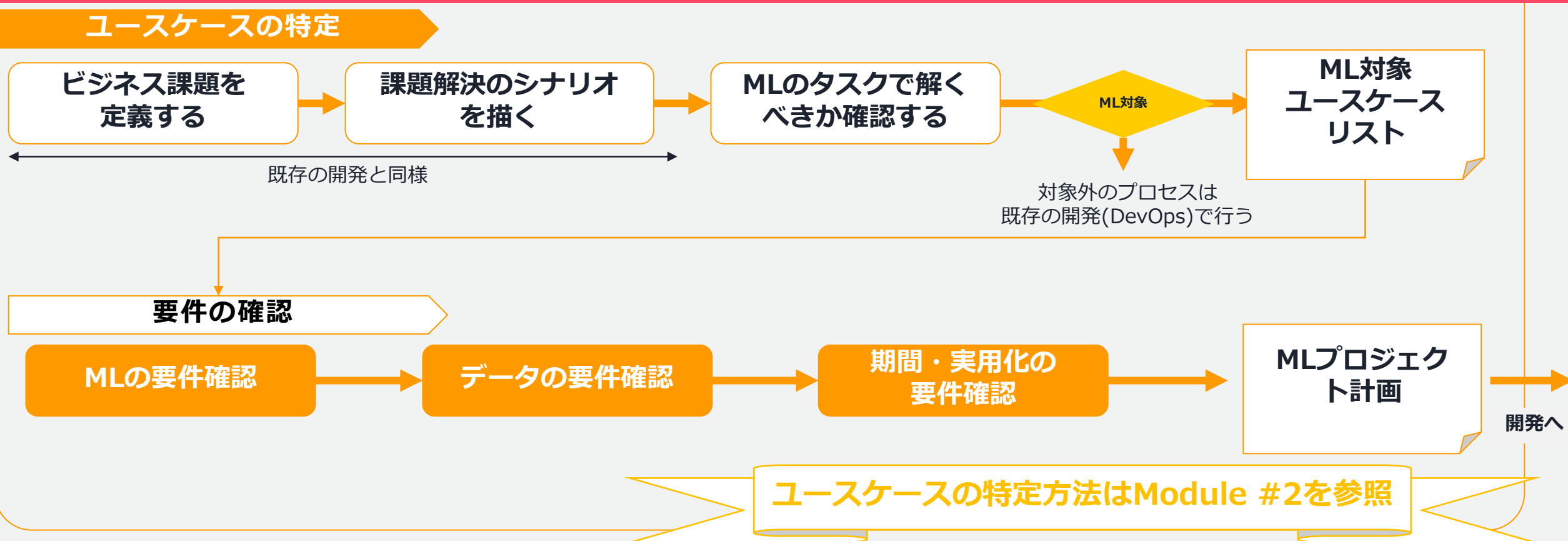


## 組織文化の変革

貴組織のニーズに合わせたMLチームをどのように作れるかを考える。

# 戦略1: 適切な問題の特定

**ML活用のも멘タム**をつけるには、ユースケースの中でも**経営層の賛同**が得られ**早く結果が出るもの**へフォーカスすることが重要



# 戦略2: クラウドの力を活用



## MLを簡単に

ワンクリックで学習・デプロイ、ビルトインアルゴリズム、自動チューニング



## スケーラブル

強力なGPUへのオンデマンドアクセス、オートスケーリング、分散学習



## コストパフォーマンス

トータルコスト(TCO)の削減、利用分だけ支払う



## セキュア

完全なセキュリティ機能: ネットワーク隔離、アクセス制御、暗号化、コンプライアンス

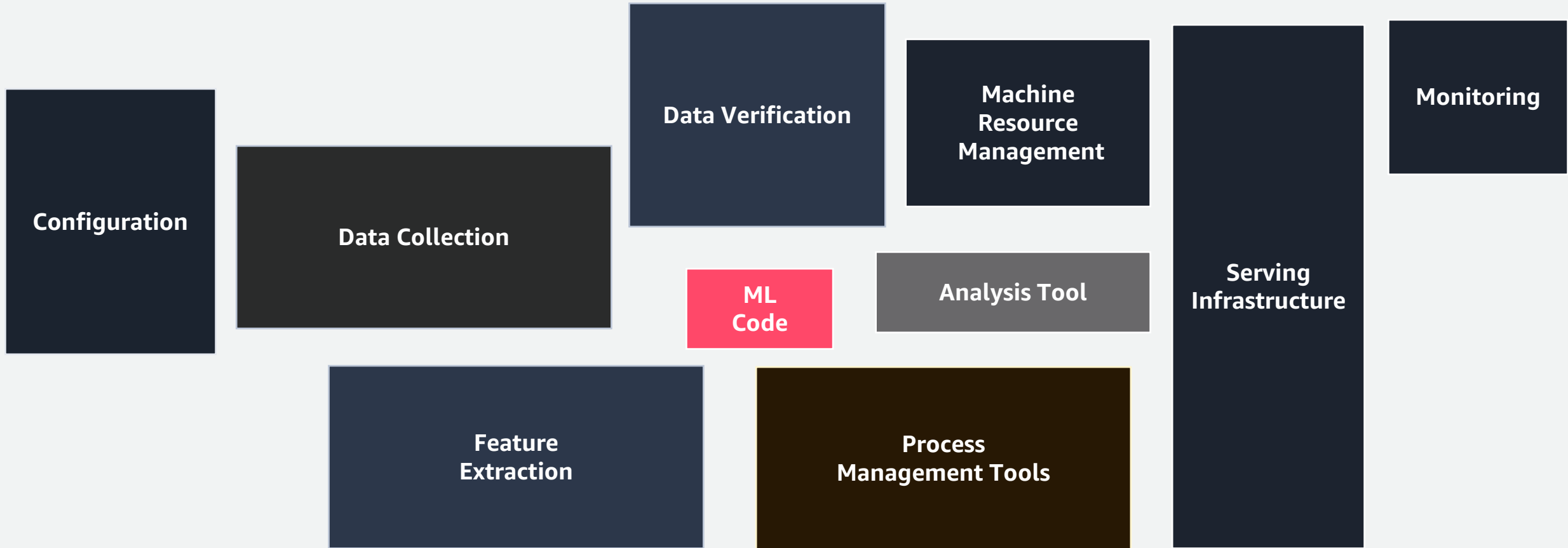


## DevOps対応

フルマネージドMLOps、Kubernetesでの稼働オプション



## 戦略2: クラウドの力を活用

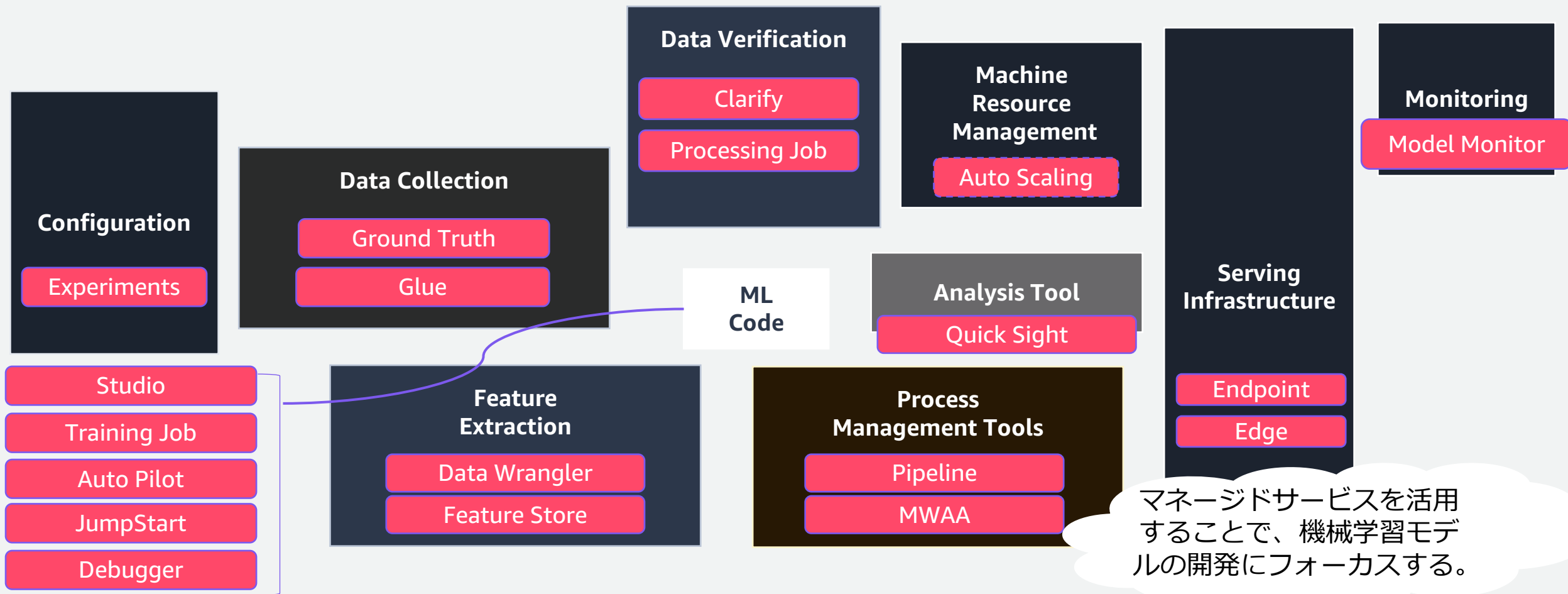


***“Only a small fraction of real-world ML systems is composed of the ML code”***

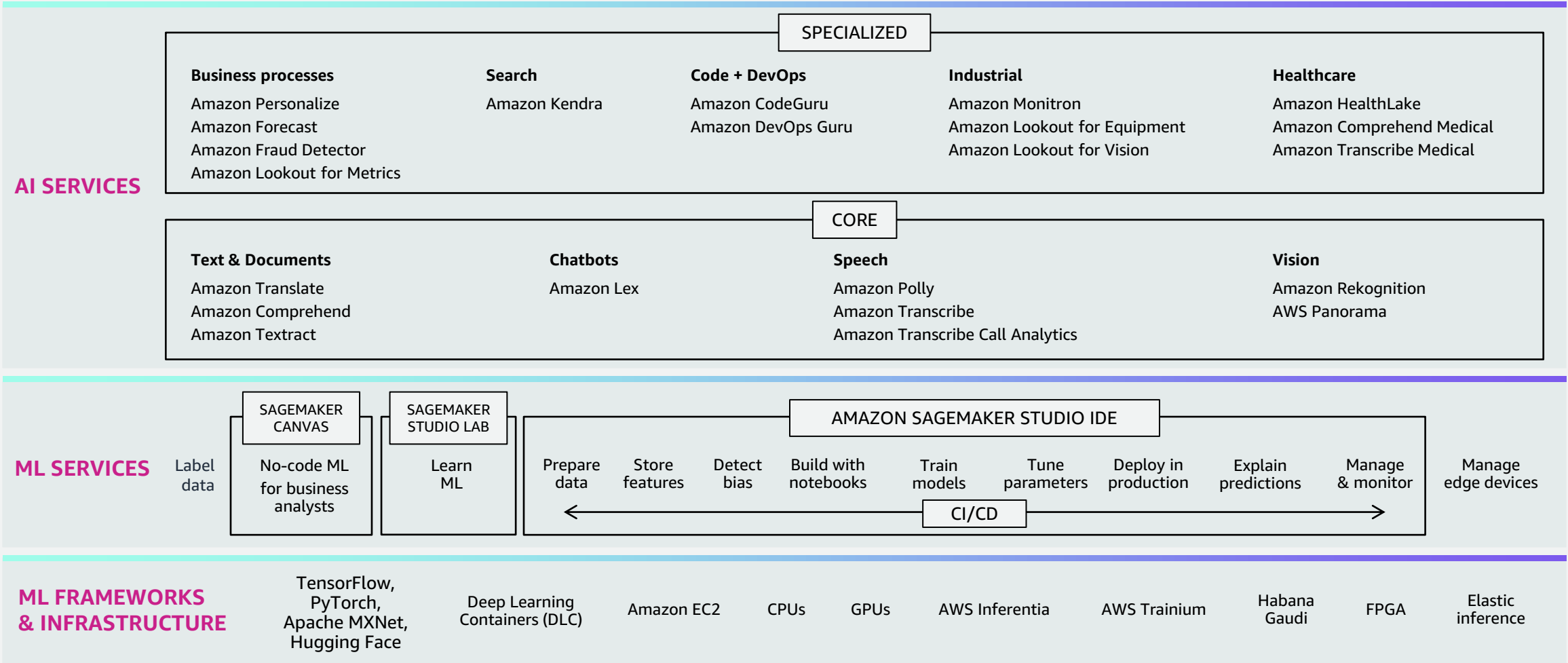
source: Hidden Technical Debt in Machine Learning Systems [D. Sculley, & al.] – 2015

<https://papers.nips.cc/paper/5656-hidden-technical-debt-in-machine-learning-systems.pdf>

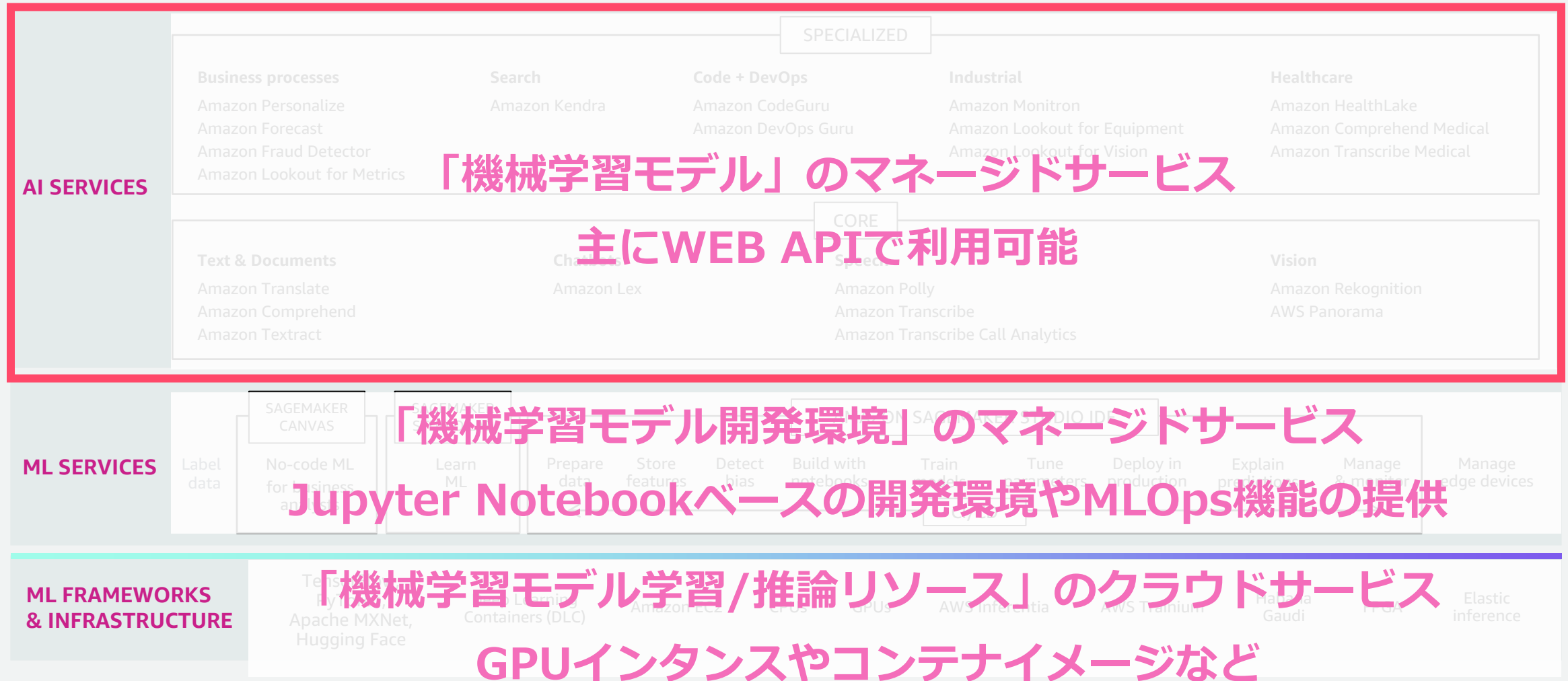
# SageMakerはML Codeに集中したい開発者のために、機械学習の技術的負債を解消するマネージドサービスを提供します



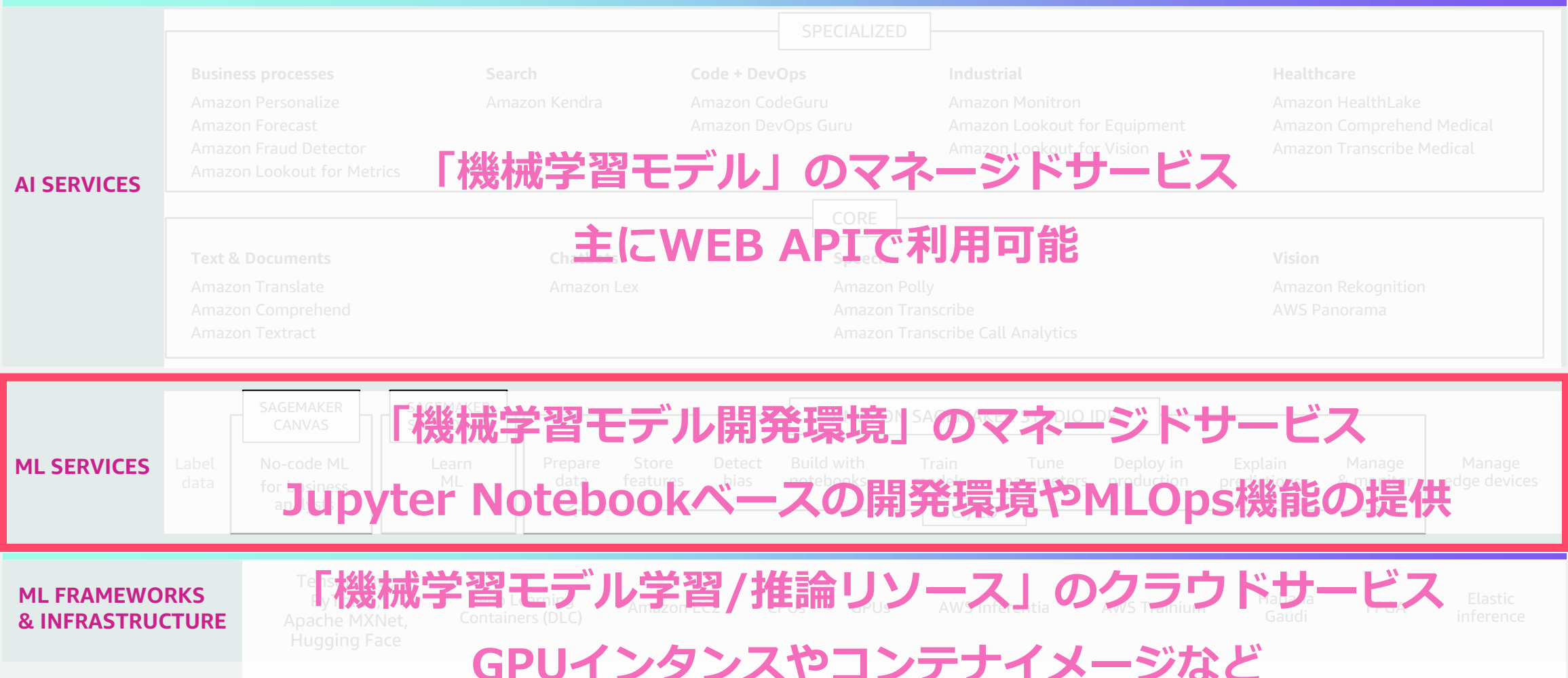
# 戦略2: クラウドの力を活用



## 戦略2: クラウドの力を活用



# 戦略2: クラウドの力を活用



# 戦略2: クラウドの力を活用

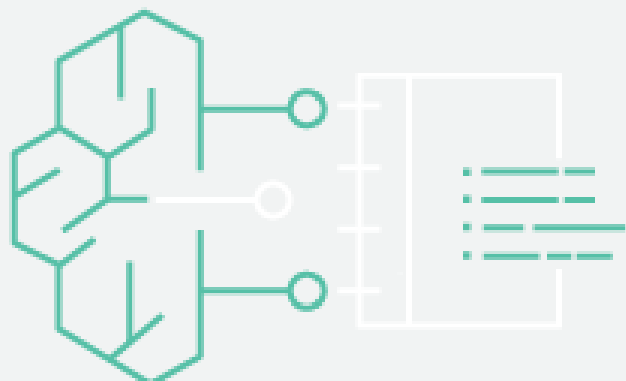
機械学習を学び、検証し、試作する

Learning

Experimenting

Prototyping

SageMaker Studio Lab



無料・メールアドレスのみで利用可

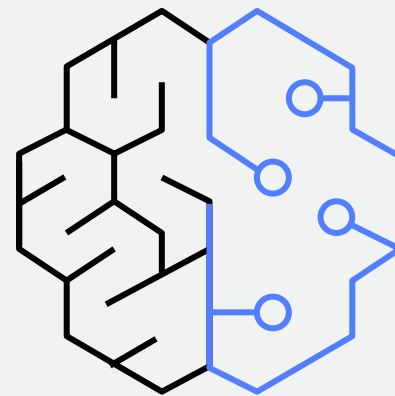
機械学習モデルを構築・運用・スケールする

Building

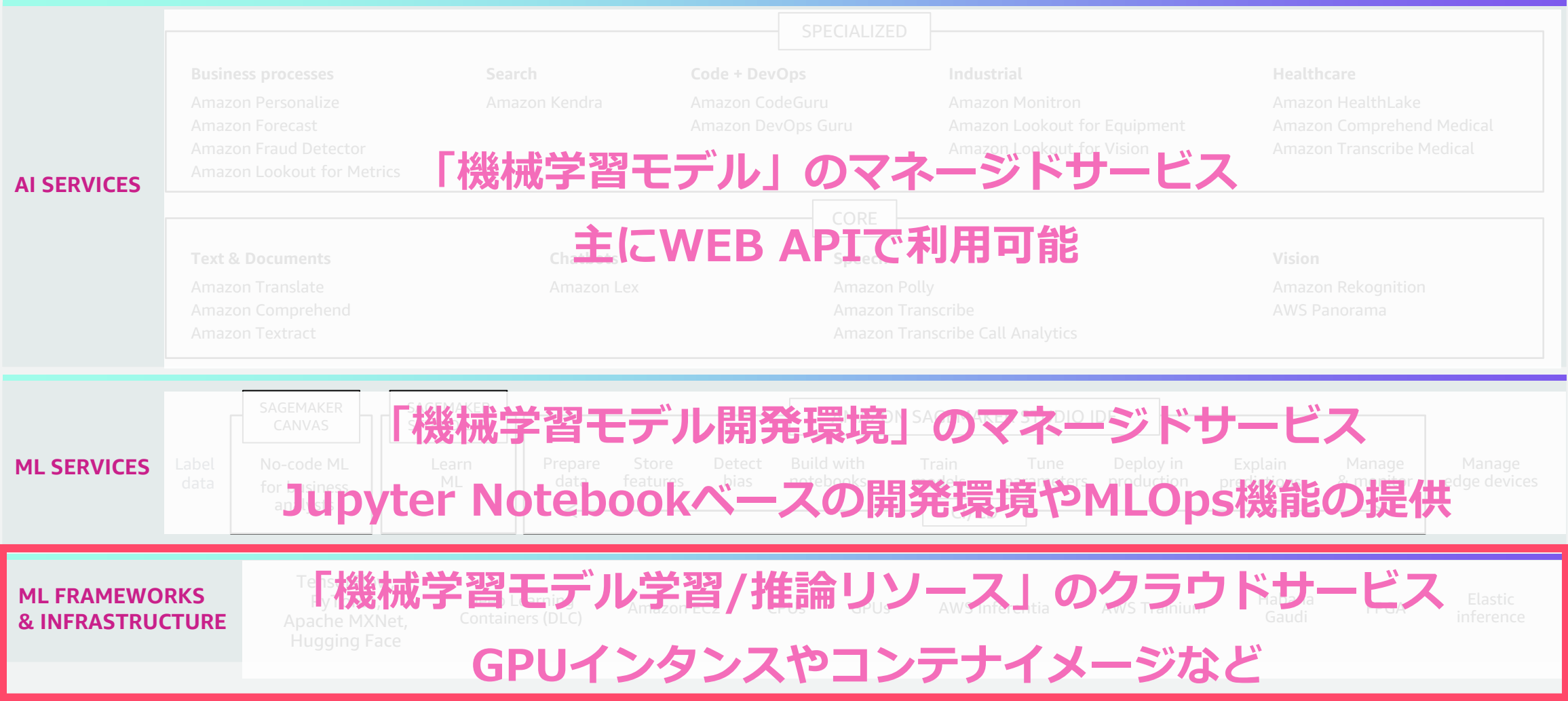
Deploying

Scaling

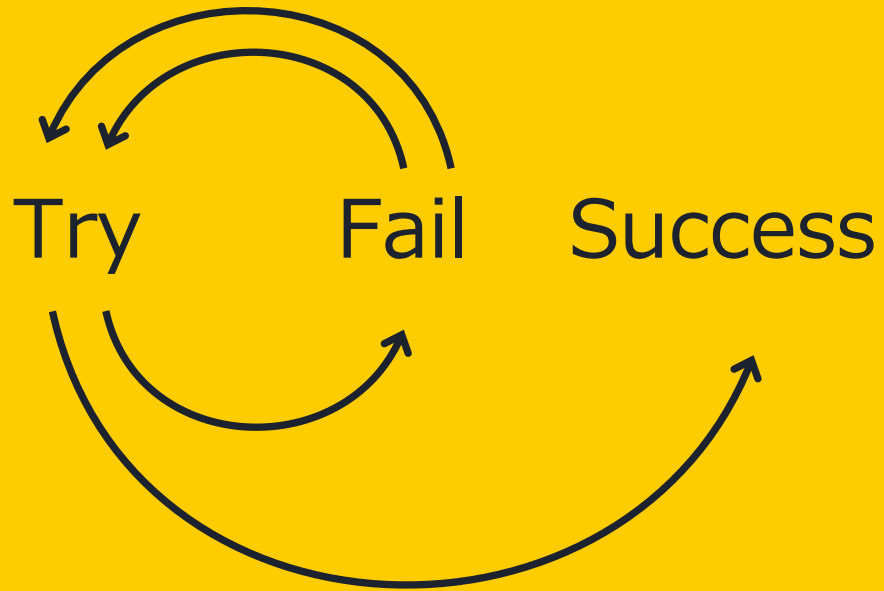
SageMaker



# 戦略2: クラウドの力を活用



## 戦略3: 前向きな失敗の奨励



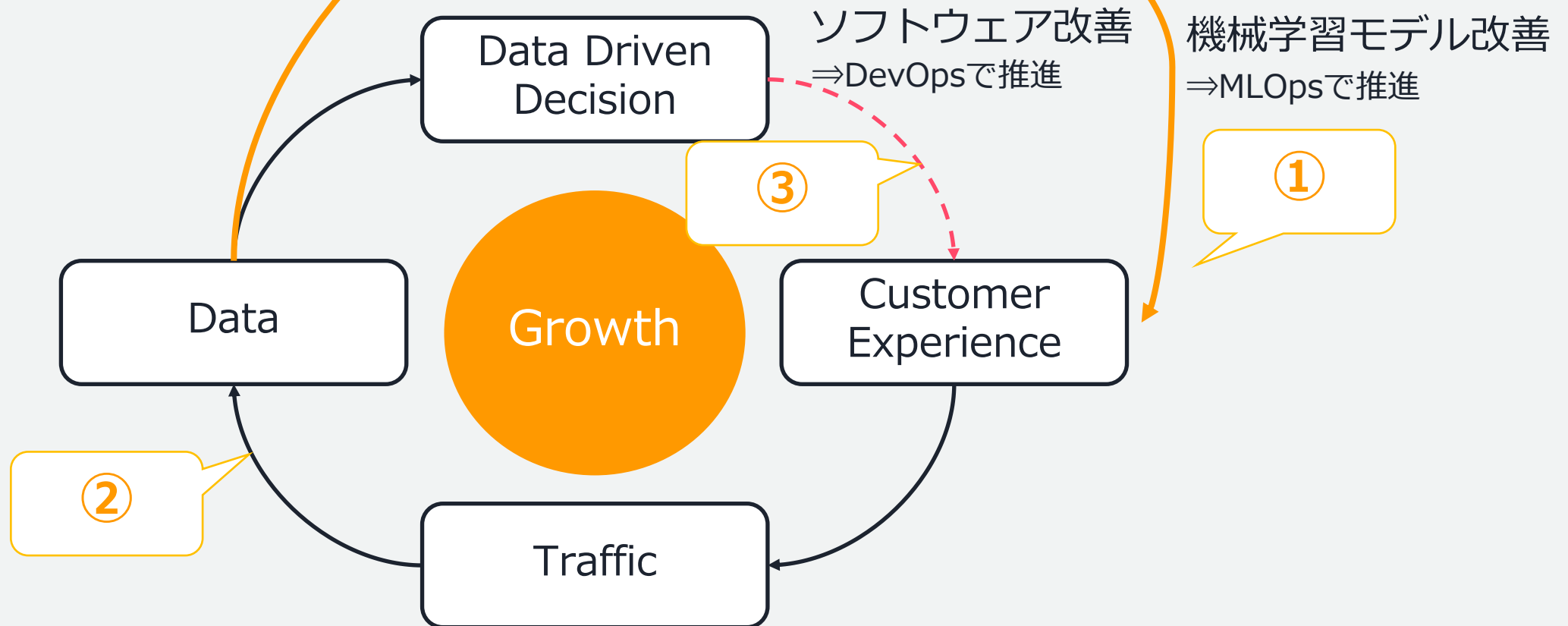
MLは反復が必要なプロセス

試行錯誤の繰り返しから学び、  
失敗への耐性を身につける

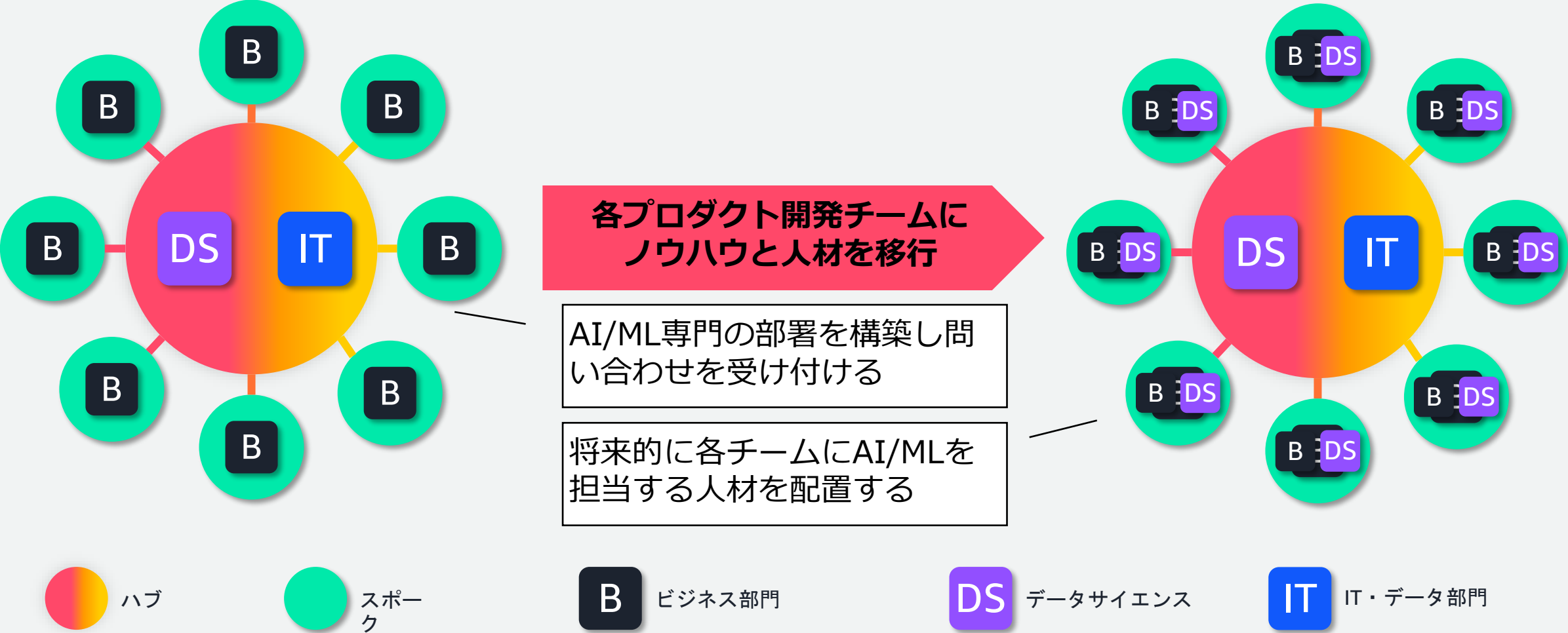


## 戦略4: POCを越えるスケーラビリティ

- ①機械学習がCustomer Experienceを改善しユーザー数が増える。
- ②増えたデータが学習データの源泉となる。
- ③モデルの改善速度と精度が、ソフトウェアによる改善を上回る。

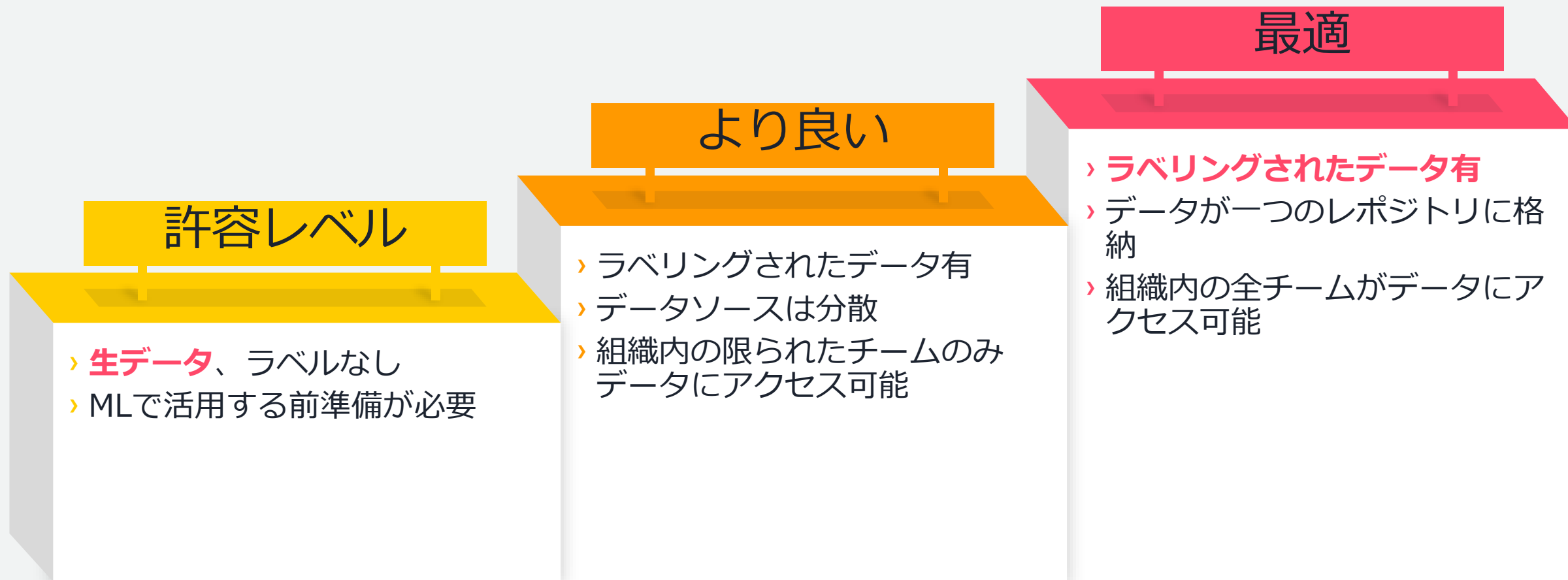


# 戦略5: カルチャーシフトのための組織作り

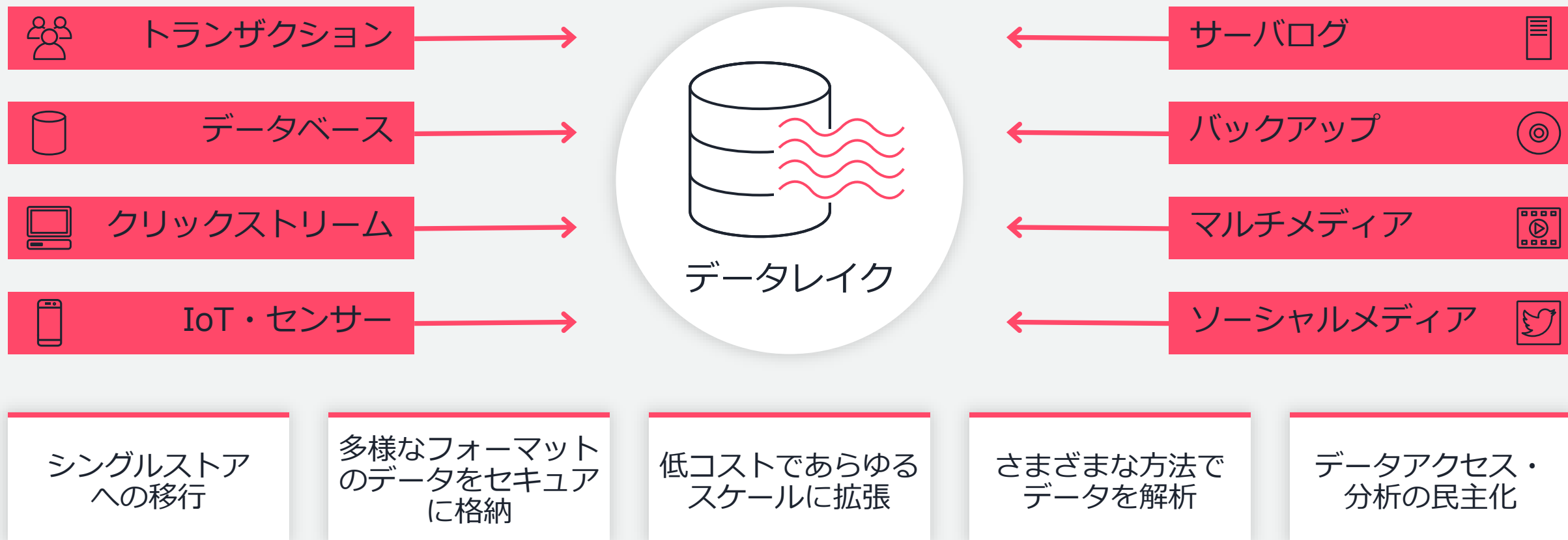


1. 堅牢なAI戦略
2. データ戦略
3. 組織内のコラボレーション
4. 推進のためのリーダーシップ

# データ戦略の推進：データの現状評価



# データレイクを真実の情報源として活用



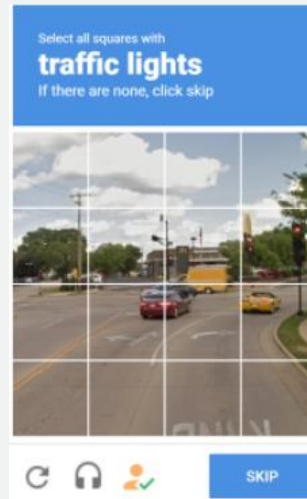
# PoCの段階からデータの収集・活用プランを設計することで最適なフェーズへいち早く移行する。



## PoCの段階からデータの収集・活用プランを作る

### Module2の復習

- ①機械学習によるCustomer Experienceの改善が進むほど、収益増につながる仕組みを作れている(スケーラビリティ/パーソナライゼーション)。
- ②サービスの機能や運用を通じ、学習データが半自動的に蓄積される仕組みが構築されている(即時的なレスポンス)。
- ③短期的に費用対効果の高いルールベースによる精度改善が、限界に達しつつあること(複雑なロジック)。



👍 いいね!

### 【パネル選択による認証】

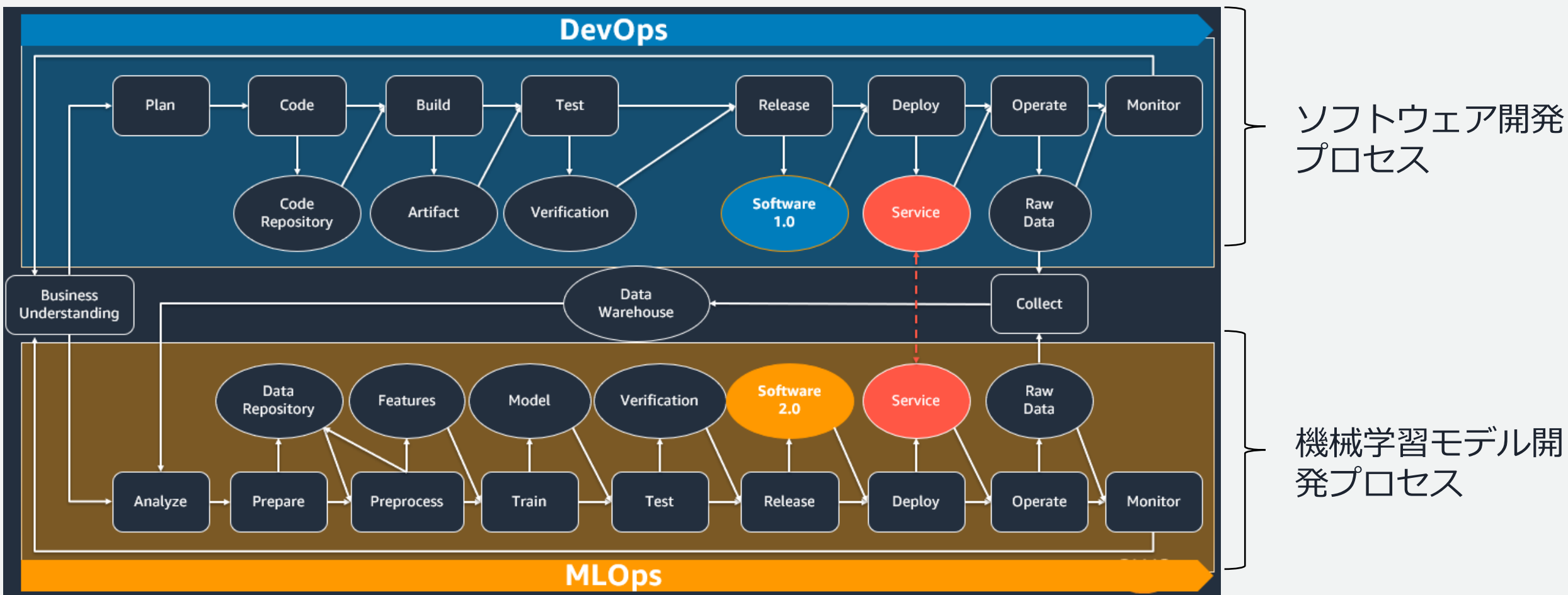
ラベルデータを生成する仕組みをアプリケーションの機能にすることでデータ作成のコストを削減できる。

### 【いいねボタン】

機械学習による提案に対し人間がフィードバックを行う機能を持つことで再学習のタイミングを検知できる。

1. 堅牢なAI戦略
2. データ戦略
3. 組織内のコラボレーション
4. 推進のためのリーダーシップ

# プロダクトに機械学習を組み込む場合の開発プロセス

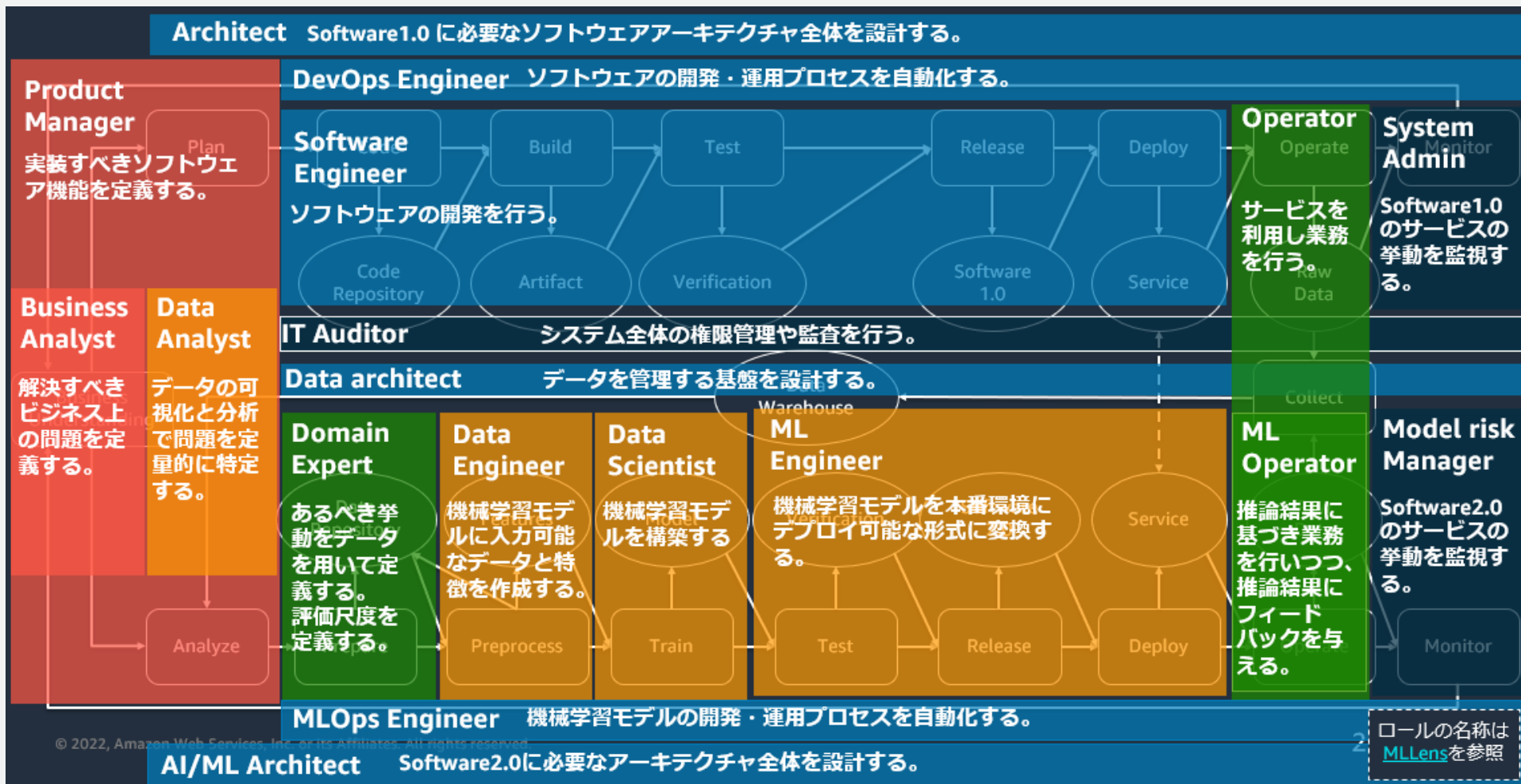


<https://speakerdeck.com/icoxfog417/mlopsfalsekoremadetokorekara>





# プロダクトに機械学習を組み込むには18の役割がかかわる



ソフトウェア開発  
チーム

機械学習モデル開発  
チーム

# パイロット機械学習モデル開発チームの一例



## プロダクトマネージャー

### 役割:

MLプロジェクトがビジネス・テクノロジー双方の要件を満たし、期限内にデリバリーされることを確認

### スキル:

プロジェクトマネジメント、リーダーシップ、MLアルゴリズムの概要の理解



## データサイエンティスト

### 役割:

MLモデルの構築

### スキル:

数学、統計、MLアルゴリズム、データ処理



## MLエンジニア

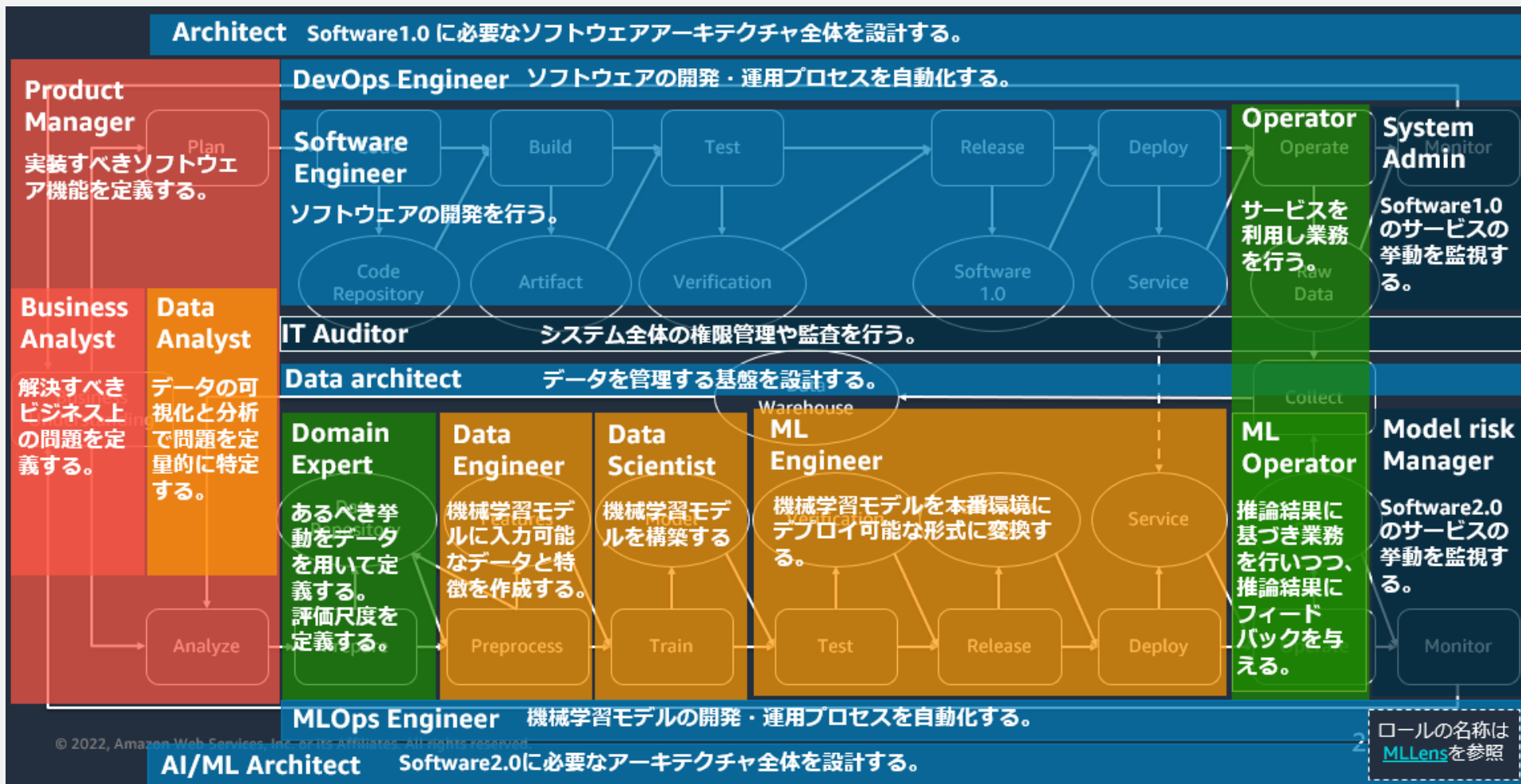
### 役割:

MLモデルのプロダクション化

### スキル:

(MLアルゴリズム)、データパイプラインツール、アーキテクチャ設計、ソフトウェア開発

# データサイエンティスト・MLエンジニアの参画



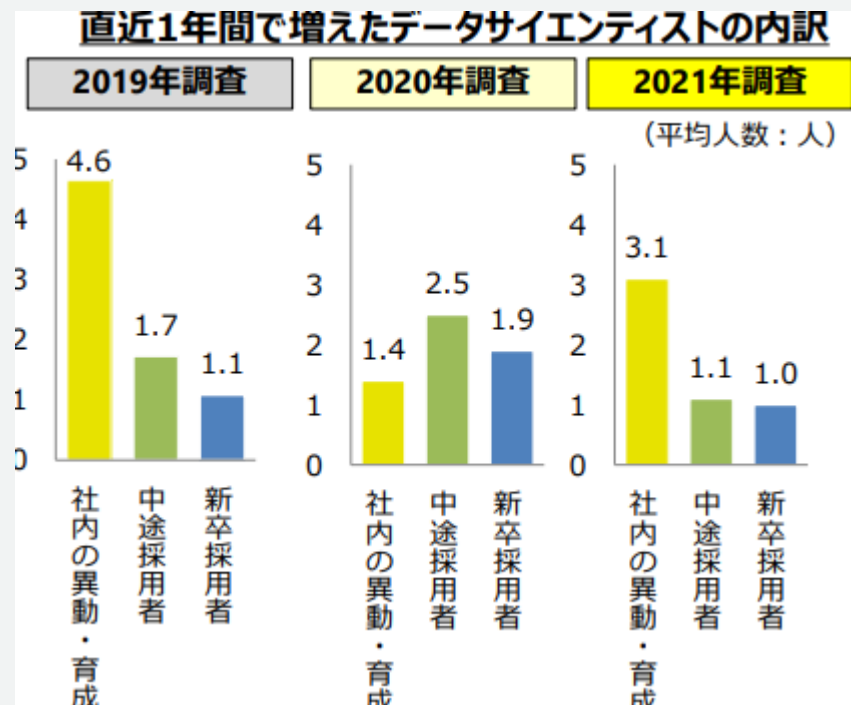
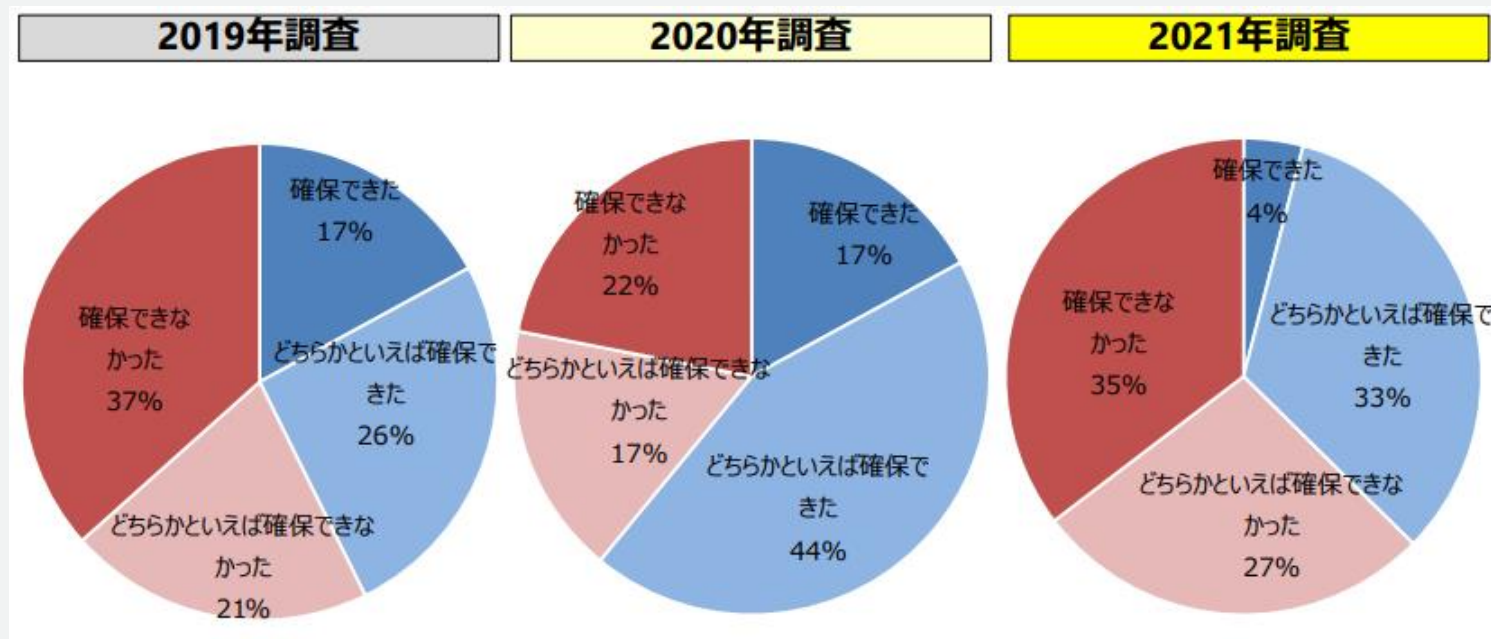
ソフトウェア開発  
チーム

スキル獲得し移動

機械学習モデル開発  
チーム

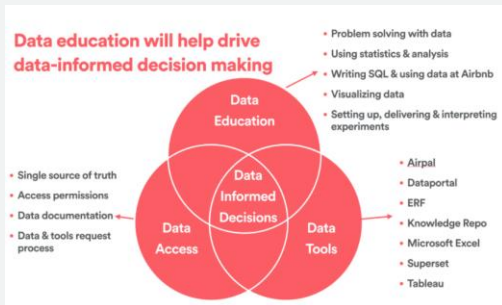
# データサイエンティストのスキルセットの典型例

データサイエンティスト協会の2021年の調査結果では、62%が目標採用人数を確保できないと回答。社内での異動・育成で補っている。



出典:データサイエンティストの採用に関するアンケート (2022年3月31日)

# 機械学習のトレーニング: 先進的企業の事例



画像引用: リンク先記事より

## [How Airbnb Democratizes Data Science With Data University](#)

データを「顧客の声」と定義し、全社員が顧客の声に基づいた意思決定ができるよう社内でのデータサイエンス教育を開始。半年で500名が受講。社内のデータを使用してトレーニングするため、30名ほどの社員がボランティアで講師を行っている。

## [Scaling Machine Learning at Uber with Michelangelo](#)

技術系の採用者全員に、1週目に機械学習とUberの機械学習基盤であるMichelangeloの講義を実施。その後も、Michelangeloのアップデートに関する講習やオフィスアワーなどを実施し利用を支援。



画像引用: リンク先記事より

Airbnbは2017年、Uberは2018年の記事。Airbnbは2017時点でFacebookとDropboxにデータ活用で後れを取っていると述べているので、データ活用の先進企業はかなり先をいっている。

# AWSが提供する機械学習のトレーニング

社内でのテクニカルスキル育成に投資しよう

ハンズオン学習

トレーニング





# データサイエンティスト・MLエンジニアの参画

## Part 1

- › 技術チーム全体（サイエンティストだけでなく）がMLを利用できるように勉強
- › 非技術系リーダーにもMLに精通するよう促す
- › トレーニング、トレーニング、トレーニング...

## Part 2

- › **できればデータサイエンティストを採用**
- › **社外のチーム（例：Amazon ML Solutions Lab、Prototyping Team）との協業によりアナリティクス/MLケーパビリティを構築**

1. 堅牢なAI戦略
2. データ戦略
3. 組織内のコラボレーション
4. 推進のためのリーダーシップ



# カギとなる質問



## 全てのチーム向け

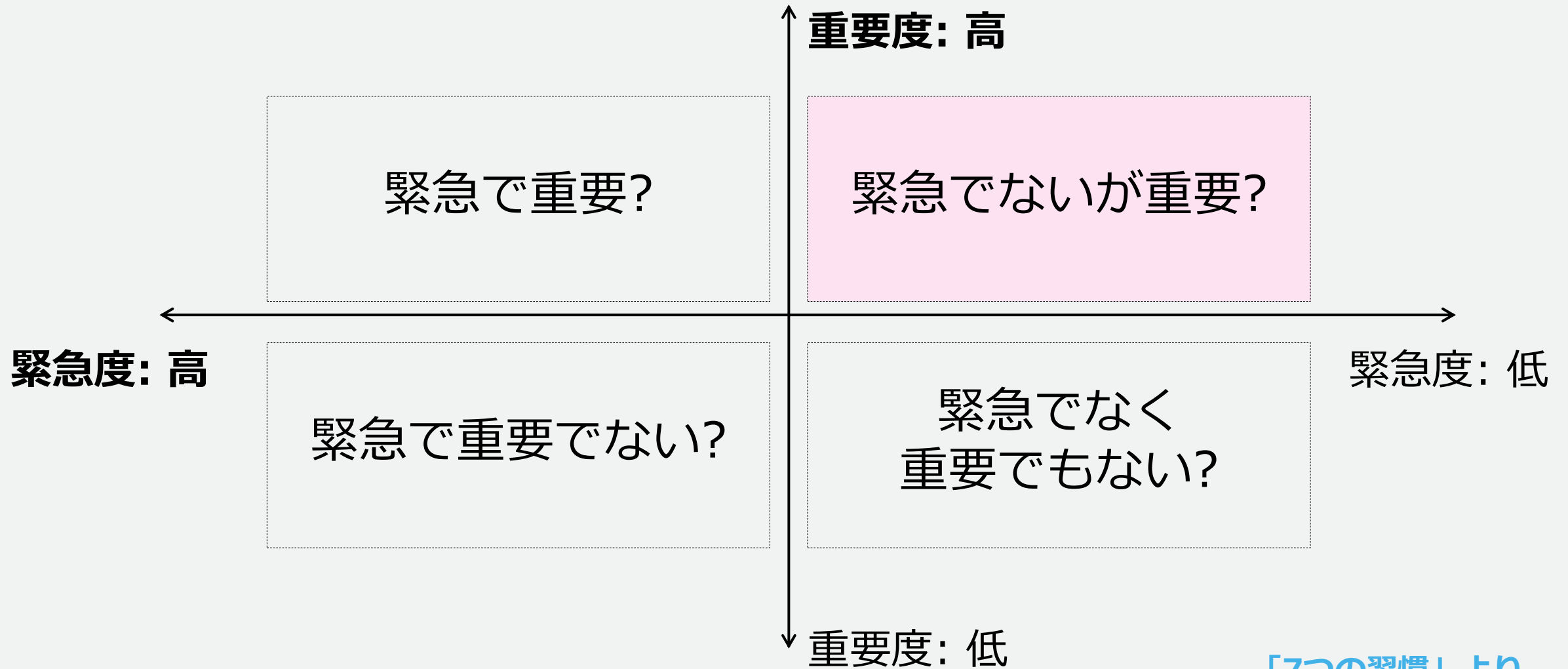
研究、ソフト開発、人事チームなども含めて  
この質問を尋ねる

「どのように機械学習を使いますか」

- › この質問への回答のために、各チームがビジネス・技術のエキスパートと協業することとなる

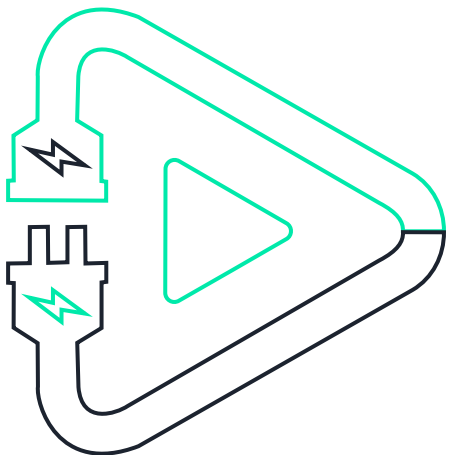
「わかりません」「使いません」  
という回答は受け入れない

# あなたのプロダクトにとって機械学習は？

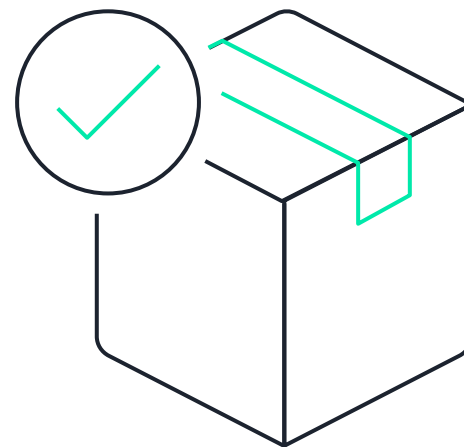


[「7つの習慣」より](#)

# よくある間違い



AIを即効性のあるプラグ&プレイ技術としてとらえる



AIの活用範囲を狭く考えすぎる

# リーダーシップのシミュレーション



考えてみましょう

あなたは、これからMLジャーニーを始めよう  
としている小売業の会社のCXOです

既にアナリティクスのパイプラインを構築し  
ていますが、MLを活用したさまざまなユース  
ケースを実現したいと考えています

**どのように始めますか？**

# MLジャーニーの開始

## ML Journey

Learning

Experimenting

Prototyping

Building

Deploying

Scaling



興味はあるが、  
実経験が浅い

### 堅牢なAI戦略

- › **重要度の高い**プロジェクトからPoCを始める
- › これらのプロジェクトを元にモメンタムを作る
- › AIへのシフトを始める理由を説明する

### データ戦略

- › 既にあるデータを探索し、ドキュメント化する
- › そのデータがAIプロジェクトに使えるかを文書で記す
- › **データパイプラインがない場合**、構築のためのタイムラインを準備する

### 組織内コラボレーション

- › **各チームにデータの探索と実験を許可する**
- › 社外リソースと協業する
- › マネージドサービスを活用する

# MLジャーニーの開始

## ML Journey

Learning

Experimenting

Prototyping

Building

Deploying

Scaling



少数の実用化  
実績

### 堅牢なAI戦略

- › 組織全体でAIケーパビリティを構築する
- › **PoCからプロダクションに移行する戦略を立てる**
- › チームを横断したコラボレーションを奨励する

### データ戦略

- › **複数のデータソースを一つの「真実の情報源」（データレイク、データウェアハウスなど）に統合する戦略を立てる**
- › 組織全体のデータリテラシーを高める

### 組織内コラボレーション

- › **社内AIチームを結成する**
- › 既存のワークフォースを採用し、オンラインや対面型のコースでトレーニングする

# MLジャーニーの開始

## ML Journey

Learning

Experimenting

Prototyping

Building

Deploying

Scaling



複数の実用化実績により  
ビジネスインパクトを創出

### 堅牢なAI戦略

- › **ビジネスの中核に長期的なAI戦略を組み込む**
- › 長期ビジョンを維持するためのAIリーダーシップ（支援体制）を作る
- › 人材採用のパイプラインを作る

### データ戦略

- › 活用可能な**外部データを取り込み**、自社データの質と有効性を高める

### 組織内コラボレーション

- › **さまざまな社内チームと協業できるAI組織を作る**
- › データに対する探索と実験を継続する



# 重要ポイント

## チーム横断コラボレーションの改善

AIは、多様なスキルと観点を持った複数のチームが機能横断的に取り組んだ際に、最も大きなインパクトを創出する





# ナレッジチェック: AI導入の支援

以下のうち、ML導入促進の初期段階において、リーダーがとるべきではないアクションはどれ？

- a) 実験と失敗の許容
- b) データの完全性およびアクセス性の確保
- c) 完全なMLインフラストラクチャの構築
- d) 最適なユースケースを特定するための素早いPoC作り
- e) 学際的なコラボレーションの促進

# ナレッジチェック: AI導入の支援

以下のうち、ML導入促進の初期段階において、リーダーがとるべきではないアクションはどれ？

- a) 実験と失敗の許容
- b) データの完全性およびアクセス性の確保
- c) 完全なMLインフラストラクチャの構築
- d) 最適なユースケースを特定するための素早いPoC作り
- e) 学際的なコラボレーションの促進

# 本日のまとめ



# Module #1: MLの入門

## MLの入門

MLの可能性

AI/MLとは

MLの活用事例

MLの技術的制約

### › MLの特性

› AI/MLの違い

› 従来型ソフトウェアとMLソフトウェアの違い

### › MLのタスク

› 回帰・分類・ランキング・クラスタリング・  
レコメンド・クラスタリング・異常検知

› ビジネスにMLがインパクトを与えたユースケース

### › MLを活用するために注意すべき点

# Module #2: MLプロジェクトの計画

## MLプロジェクトの 計画

ユースケースの特定

ML拡大の要件

データの要件

期間・実用化の要件

- › **MLが適用可能なユースケースの特定方法**
- › MLを適用するための要件
  - › MLの勝ちパターンでスケールするか
  - › データは利用可能か
  - › 実用化するための十分な期間がとれているか

# Module #3: ML活用組織へのシフト

## ML活用組織への シフト

堅牢なAI戦略

データ戦略

組織内のコラボレーション文化

推進のためのリーダーシップ

- › MLを活用できる組織に移行するためのポイント
  - › 経営層が関わる実現性の高いユースケースから着手する
  - › 自動的にラベル付きデータが蓄積する仕組みの構築
  - › ソフトウェア開発チームとのコラボレーション
- › 3つのポイントを推進するためのリーダーシップ
  - › MLの活用段階ごとに注力する点を意識



# Thank you!

This material is based on the AWS Machine Learning Embark Program,  
an immersive ML training program provided by Amazon Machine Learning Solutions Lab