

2018年7月31日『人間中心のA I 社会原則検討会議』

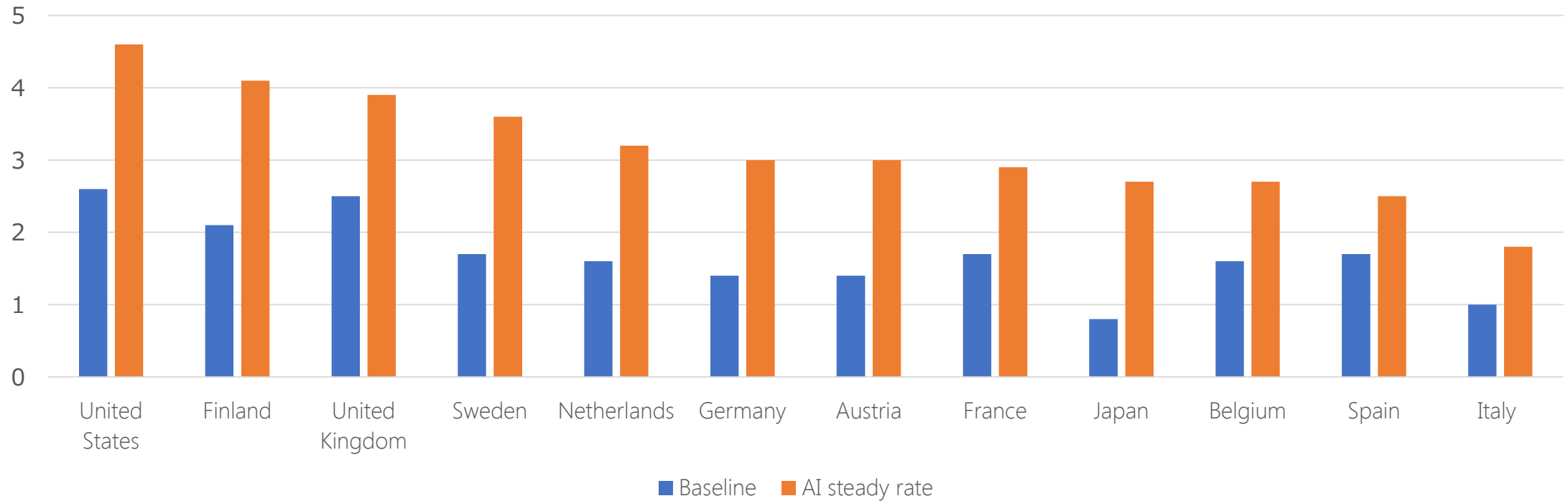
保証できるAIの実現のために

日本マイクロソフト株式会社



# AIによる経済への影響

Value creation to GDP growth from AI



# コンピューターが理解する世界

知覚



ビジョン



音声

理解

文字

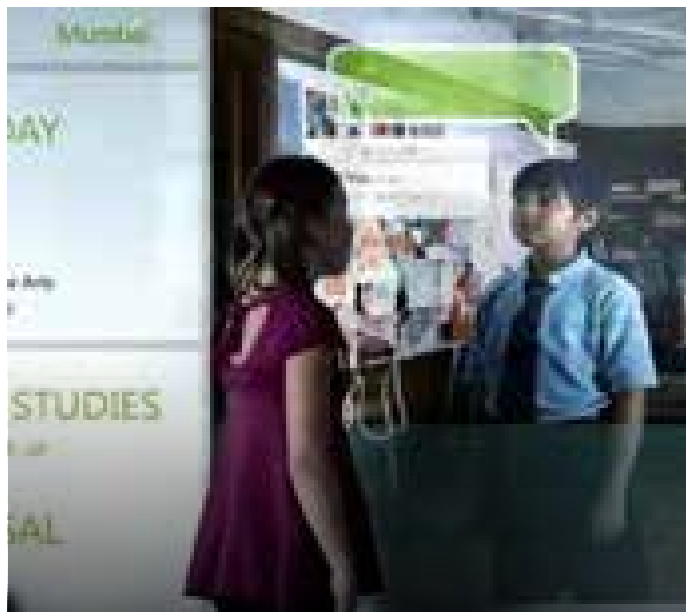
言語



知識



# AI の可能性



教育



医療



輸送

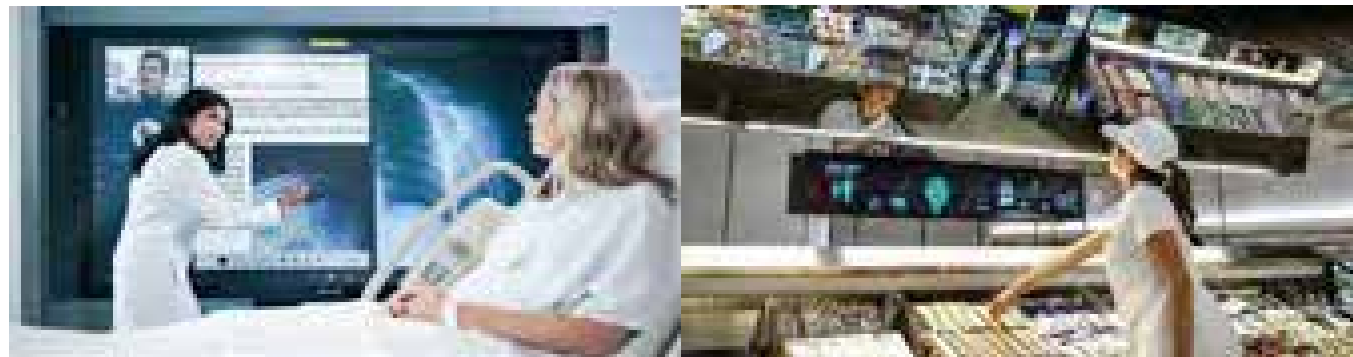


農業



I think it's a young girl throwing

マイクロソフトのビジョン:  
人の創造性を加速する  
AIのデザイン



推論



理解



相互作用

マイクロソフトの目的:

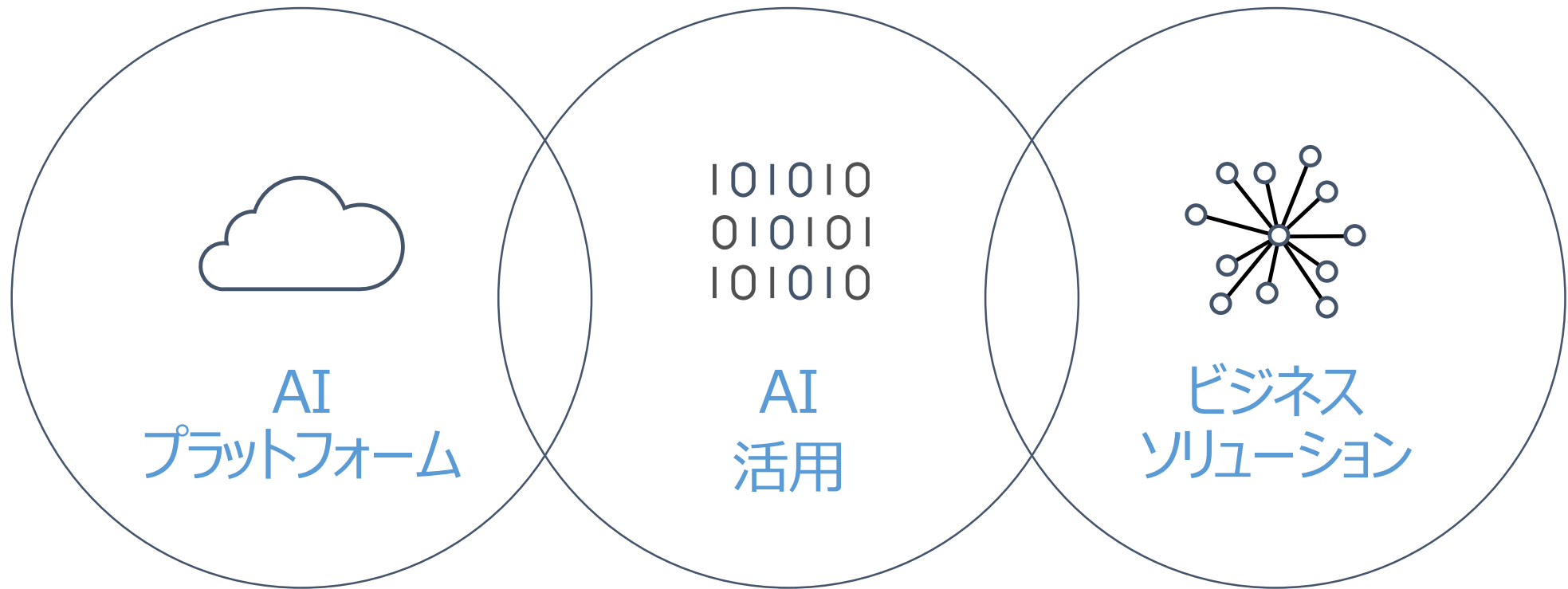
誰もが活用可能なAI



私たちはすべての人とすべての機関をエンパワーすることに力を注いでいる... 差し迫った社会と経済の問題を解決するために。

-サティヤ Nadella, CEO, マイクロソフト

# マイクロソフトと AI





# 原則、方針および法律 Responsible AI



# 信頼されるAIのために



公平性



信頼性  
& セーフティ



プライバシーと  
セキュリティ



インクルーシブ



透明性



責任

# 公平性

バイアスの導入、影響への理解

多様、且つ才能あるAI従事者の確保

バイアスを検知、排除するための分析技術の開発

人によるレビューとドメイン専門知識

## The New York Times

### Facial Recognition Is Accurate, if You're a White Guy

By STEVE LOHR FEB. 9, 2018



Gender was misidentified in up to 1 percent of lighter-skinned males in a set of 200 photos.

Gender was misidentified in up to 12 percent of darker-skinned males in a set of 210 photos.



Gender was misidentified in up to 7 percent of lighter-skinned females in a set of 200 photos.

Gender was misidentified in 35 percent of darker-skinned females in a set of 271 photos.

# 信頼性と安全性

トレーニング データの評価

広範囲にわたる検証 (および常にユーザーフィードバックを得るための仕組み)

継続的なパフォーマンス モニタリング

悪意ある攻撃を含む予期しない状況を想定したデザイン

常に人が介在する仕組み



# プライバシーとセキュリティ

既存のプライバシーに関する法律（例：[改正個人情報保護法](#)）の適用

データ収集、使用、および[管理](#)に関する[透明性](#)を担保することにより、データの選択を可能に

悪意ある利用からの[保護](#)を前提としたシステム

プライバシーおよびセキュリティの為の  
[匿名化](#)、[非識別化技術](#)の活用



# インクルーシブ

潜在的なバイアスによる意図しない排除を防ぐインクルーシブデザイン

障がいを持つ人たちの可能性の最大化

信頼性の構築

IQだけでなく、EQ



# 透明性

どのように判断されたかを理解可能に

包括的な情報の提供

潜在的なバイアス、エラー、そして意図しない結果への意識を高める



# 責任

人が責任を負うシステム運用

システムのデザインにおける規範の継続的な監視

内部監査と役割





# AI とこれからの仕事



# 働き方の変革

## オンデマンド経済

いつでも、どこでも

仕事の在り方、働き方の変化



Virtual space designer



Biohacker



Trainer for autonomous cars



IoT data analyst

# PAI Mission

開発と共有のベストプラクティス

AIへの理解を促進

議論のためのオープンプラットフォームを提供

社会的に有益なAI活用



# 倫理的枠組みの検討

広範な展開そして継続的なイノベーションのために:

- マルチステークホルダー による対話
- 信頼できる人間中心の AI のための実用的な原則
- ベストプラクティスの共有
- 研究 & 開発投資
- データ の可用性
- スキル トレーニング



より確かなAIによる  
エンパワーメント





©2013マイクロソフトコーポレーション。すべての権利予約。マイクロソフト、windows、windows Vista およびその他の製品名は、米国および/またはその他の国の登録商標および/または商標である場合があります。本書は情報提供のみを目的としており、このプレゼンテーションの日付時点でのマイクロソフトコーポレーションの現在の見解を表しています。マイクロソフトは、変化する市場状況に対応するため、マイクロソフトの一部に対するコミットメントであると解釈すべきではなく、マイクロソフトはこのプレゼンテーションの日付以降に提供された情報の正確性を保証することはできません。マイクロソフトは、このプレゼンテーションの情報に関して、明示的、黙示的、または法定の保証を行いません。