要求に応じて
AIを仕立て上げる

ディペンダブルな深層学習システムのための工学技術

何が問題？

機械学習技術を用いた人工知能（AI）システムでは、データからの学習を通して複雑な動作を実現することができます。
しかし、エンジニアが動作規則を設計していた従来ソフトウェアと異なり、**動作規則を詳細に把握し制御することが困難**です。
この点は、自動運転など安全・信頼が重要な領域でAIシステムを活用するための大きな課題となっています。

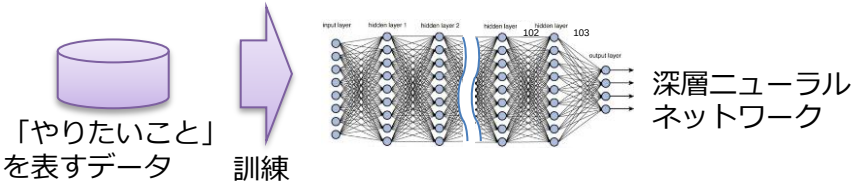
何を目指す？

従来ソフトウェア（つまりプログラム）に対しては、
「不適切な動作の発見、原因の分析、そして修正」
というテスト・デバッグの技術が発展してきました。
これらの技術をAIシステムに適合させることにより、
細やかな要求に応えるようシステムを改善し続けていくための
技術の確立を目指しています。

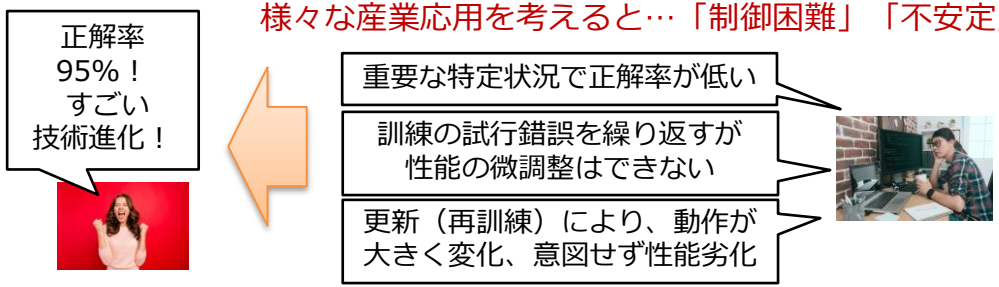
背景

機械学習、特に深層学習技術を用いたAIシステムの開発：

人間には意味の理解が困難な単純計算の膨大な集まりに対し
訓練データを基に膨大なパラメーター設定を行う（software 2.0）



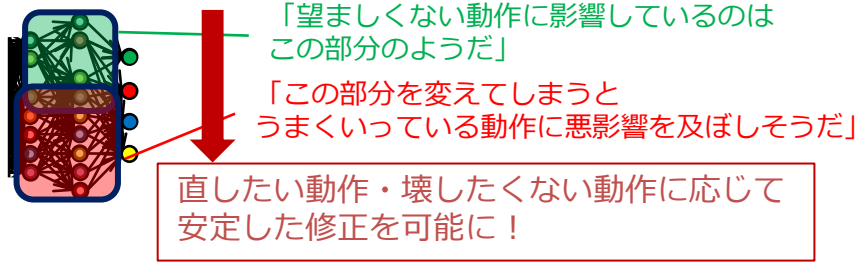
様々な産業応用を考えると…「制御困難」「不安定」



研究内容

従来：プログラム実行の成功・失敗の原因箇所を推測する技術
（Fault Localization）

AI（深層ニューラルネットワーク）においても
成功・失敗に影響している箇所を推測



JST未来社会創造事業 “Engineerable AI” (eAI) プロジェクト
6/28 キックオフシンポジウム https://engineerable.ai/