

# AI活用が研究競争力を生む時代へ ～問題解決プロセスの進化～

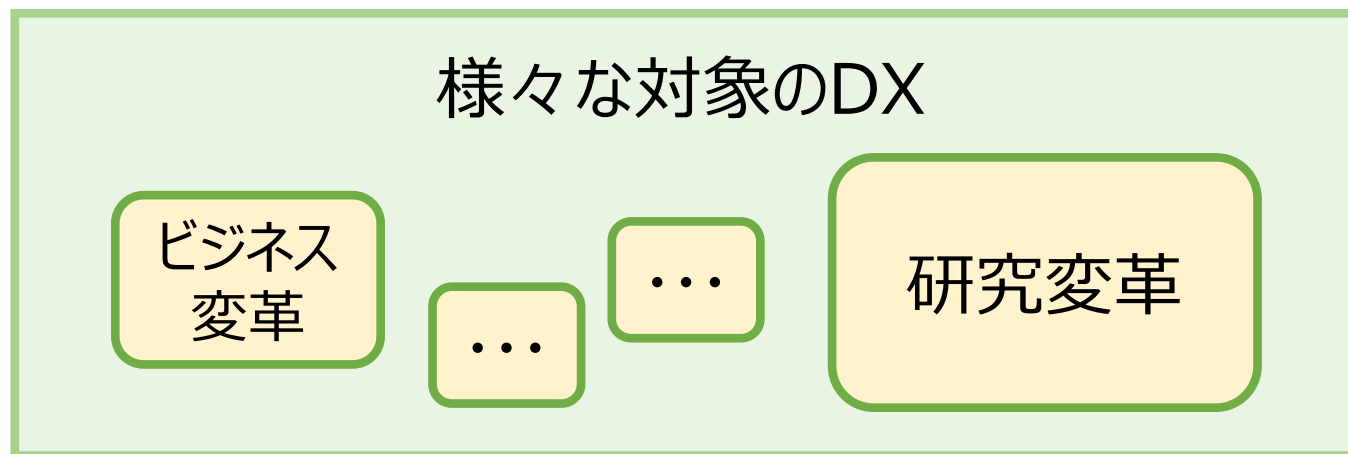
2020年9月28日～11月30日

システム・情報科学技術ユニット 福島俊一



国立研究開発法人科学技術振興機構 研究開発戦略センター  
Center for Research and Development Strategy Japan Science and Technology Agency

# 講演の構成



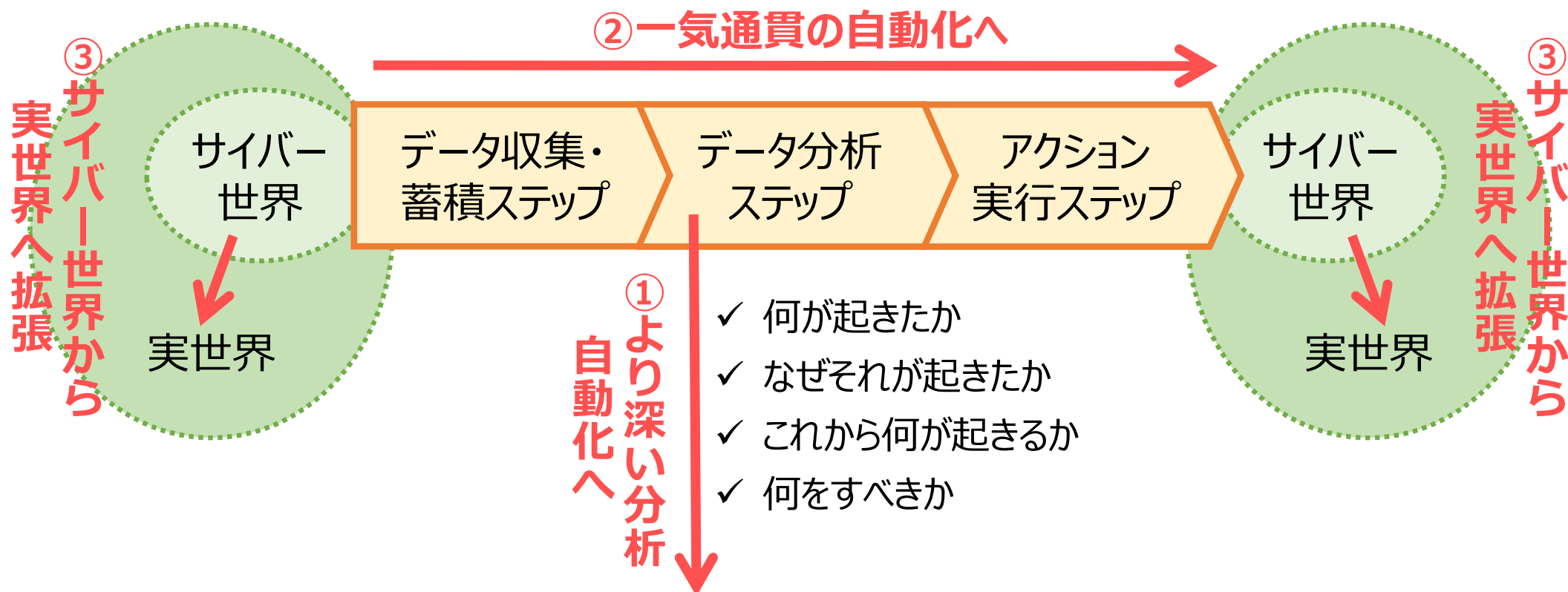
2. 講演の後半



1. 講演の前半

# データに基づく問題解決プロセスの進化

- ビッグデータとAI技術の発展・普及によって、データに基づく問題解決プロセスが広く活用されるようになった
- このプロセスは、3方向の技術的発展によって進化している



# 技術的發展① より深い分析自動化へ

- データ分析の深さは4段階あると言われる (Gartner社による分類)
- 機械学習、最適化、推論、シミュレーション等のAI関連技術が発展し、より深い分析の自動化が可能になってきた

データ分析の4段階		自動化のための技術の例
1	Descriptive: <b>記述的分析</b> 「何が起きたか」	機械学習による画像認識・異常検知、クラスタリング・可視化等
2	Diagnostic: <b>診断的分析</b> 「なぜそれが起きたか」	機械学習による判別・分類、画像・言語理解、関係分析
3	Predictive: <b>予測的分析</b> 「これから何が起きるか」	機械学習による予測、回帰分析・時系列分析
4	Prescriptive: <b>処方的分析</b> 「何をすべきか」	深層強化学習、予測型意思決定最適化、知識ベース推論、マルチエージェントシミュレーション等

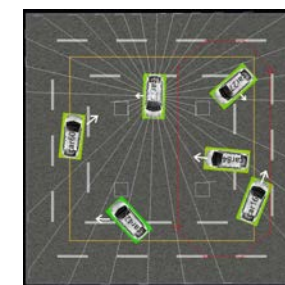


映像認識による  
群衆行動把握

[https://jpn.nec.com/press/201503/20150310\\_01.html](https://jpn.nec.com/press/201503/20150310_01.html)

異種混合学習による  
来店客数予測

<https://wisdom.nec.com/ja/events/2018071101/index.html>

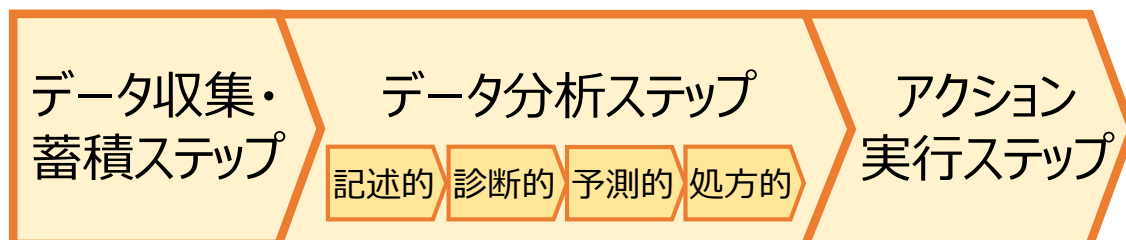


深層強化学習による車の  
運転制御

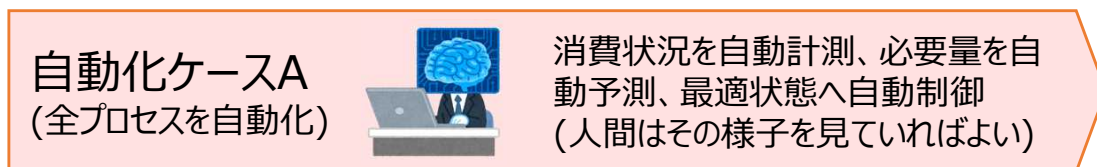
<https://tech.preferred.jp/ja/blog/ces2016/>

## 技術的发展② 一気通貫の自動化へ

- 問題解決プロセス全体にわたって自動化されるほど、人間の負荷が軽減され、ソリューションとしての価値が高まる
- これまでデータ分析に関わる特徴設計・モデル選択・ハイパーパラメータチューニング等は、人手で試行錯誤しなくてはならなかったが、その自動化技術(AutoML)の性能が高まってきた



例えば電力マネジメントの自動化



### AutoML

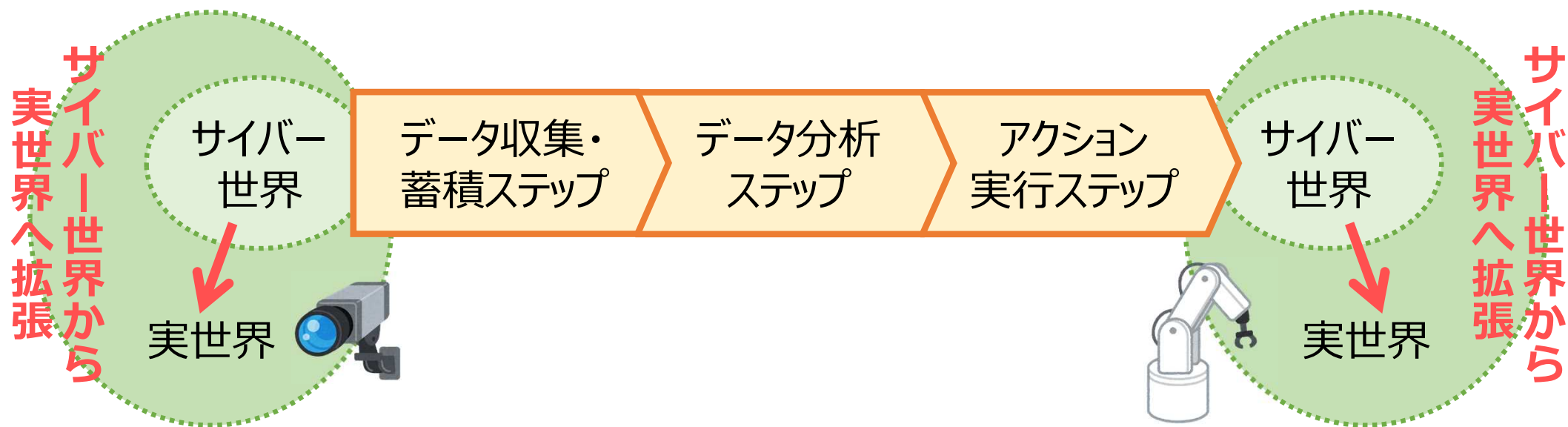
(Automated Machine Learning)

機械学習における特徴設計・モデル選択・ハイパーパラメータチューニング等を自動化、データ分析に要する期間が大幅に短縮されることに加えて、必ずしも高い専門知識を持たない人でもデータ分析を簡単に実施可能になる

代表的なAutoMLシステム

- DataRobot
- dotData
- H2O.ai
- Google AutoML 等

# 技術的发展③ サイバー世界から実世界へ拡張



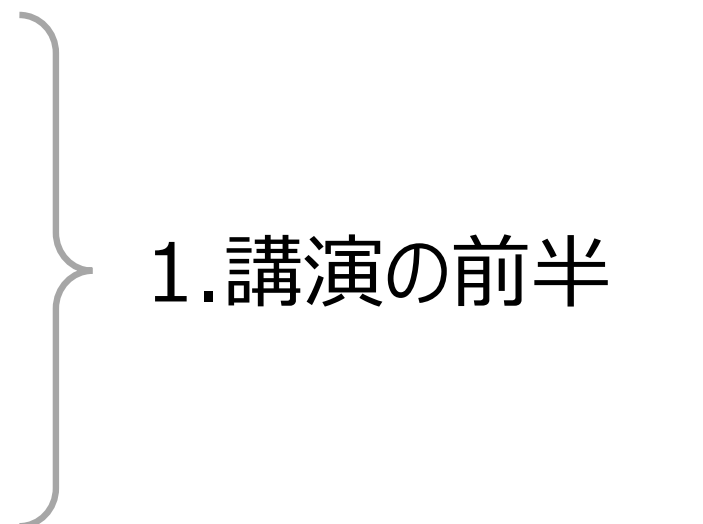
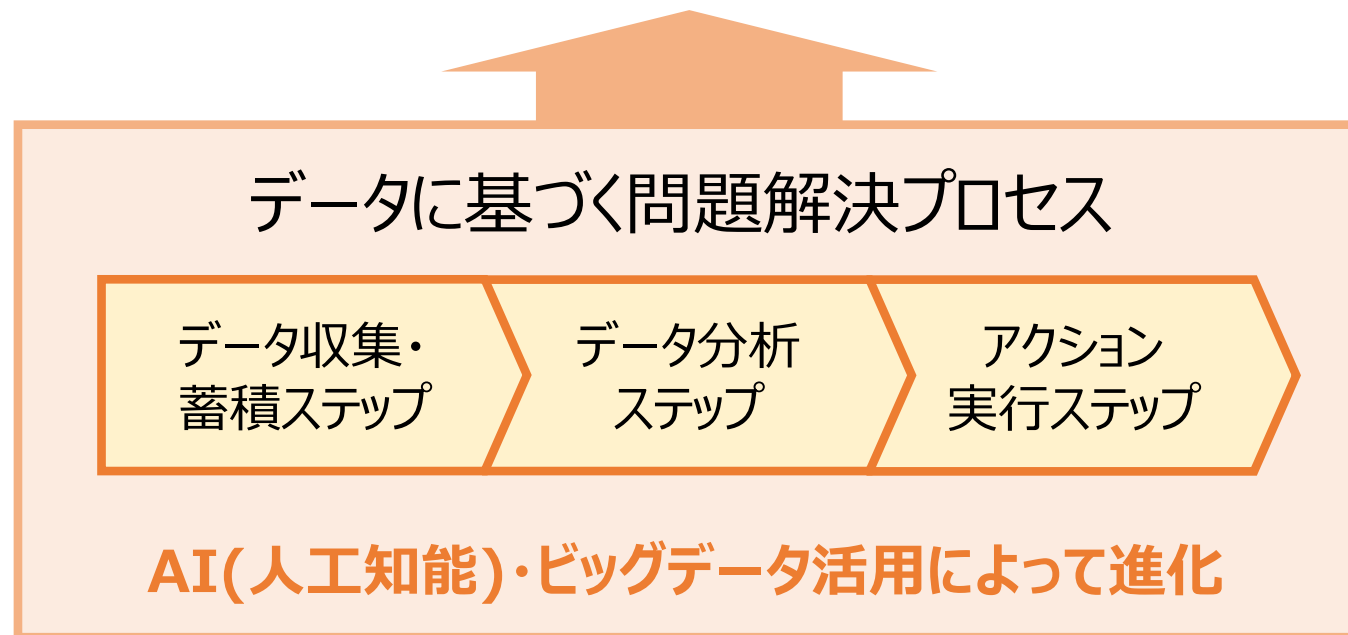
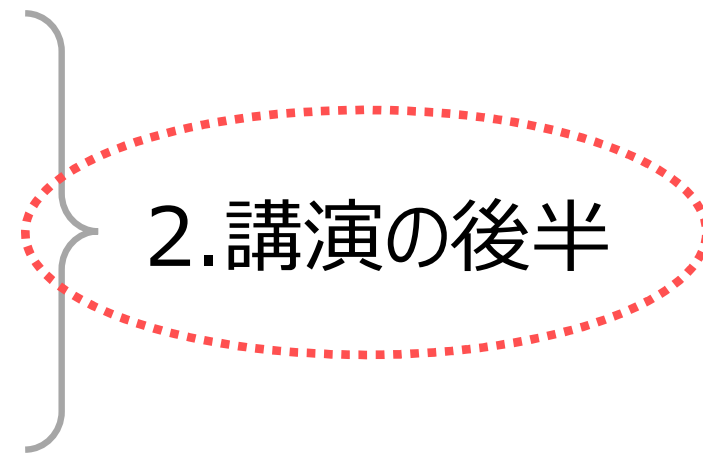
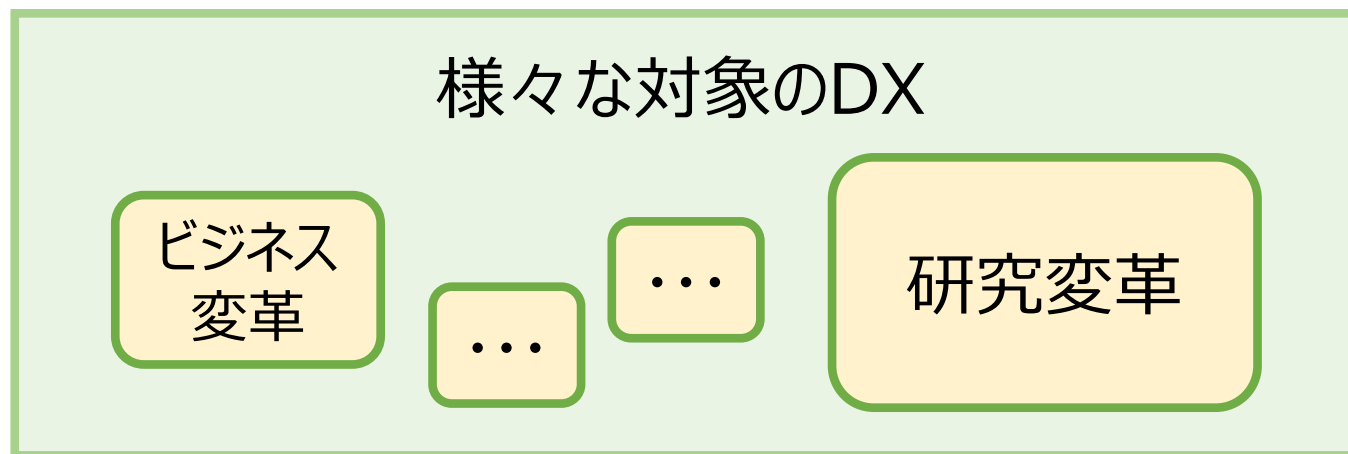
- センサー、カメラ、アクチュエーターを含む様々なIoTデバイスやロボット技術が導入され、データ収集・蓄積ステップとアクション実行ステップがサイバー世界だけでなく実世界にも広がった
- IoTデバイスやロボット技術に関して、軽量化、省エネ化、高感度化、高解像度化、スマート化等の技術改良が進められている

※この技術的发展③について、CRDSストリーミングセミナーシリーズでは「サイバーフィジカル融合を高度化するデータ取得技術～センシング・ロボット技術の活用～」にて詳しい動向の紹介があります

# 問題解決プロセスを支えるデータ基盤技術

データ基盤技術		主な注目技術
<b>データ処理基盤技術</b>	大規模なデータを高速に処理するための技術群	分散並列処理技術、圧縮データ処理技術、ストリームデータ処理技術等
<b>データ保護基盤技術</b>	対象となるデータを保護するための技術群	暗号化等のセキュリティー技術、プライバシー保護技術、データの分析と保護を両立させるプライバシー保護型データ分析技術等
<b>オープンデータ技術</b>	データを最小限の制約のみで誰でも自由に利用・加工・再配布できるようにするための技術群	様々なデータ活用を促進する共通的なデータ形式や付加的情報(メタデータ)の記述形式のデザイン、セマンティックウェブの標準化技術を用いたリンクト・オープンデータ、分野・組織をまたいだデータ連携を容易にする共通語彙・データ連携基盤等

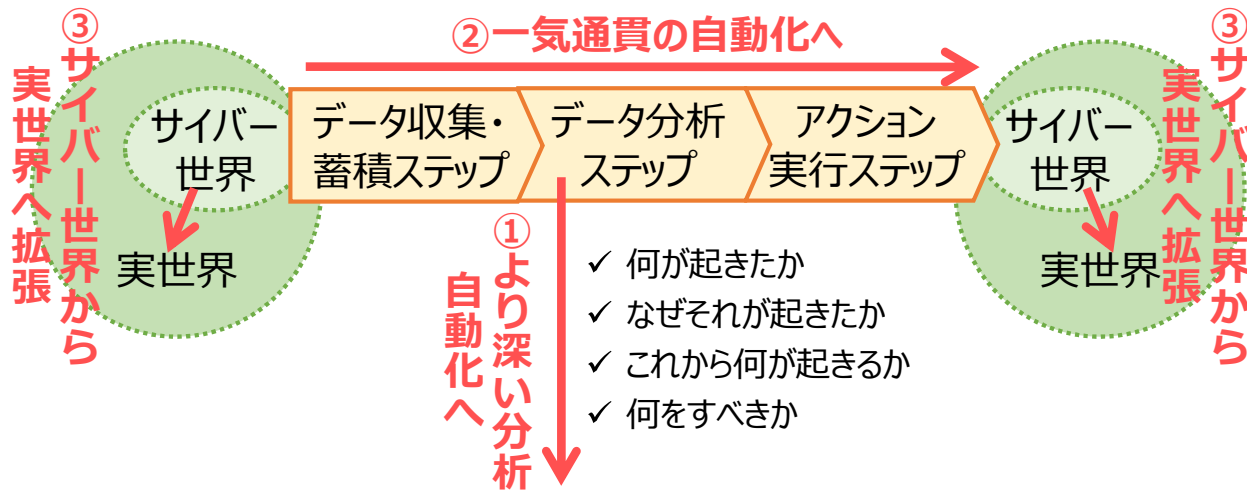
# 講演の構成





# 問題解決プロセスの進化がもたらすビジネス変革

## データに基づく問題解決プロセスの進化



## もたらされる効果

- 実世界で起きる様々な現象や活動の精緻かつリアルタイムの把握・予測
- 膨大な選択肢の網羅的な検証
- 大規模複雑タスクの自動実行等

## ビジネス変革の例

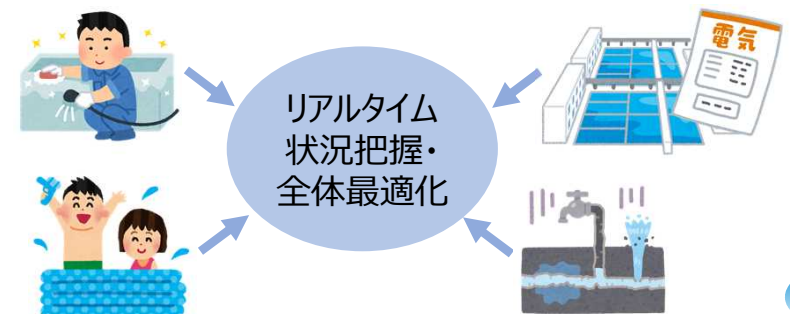
自動化によって参入障壁が低下し、従来の業界の枠にとらわれない新しいビジネスが生まれ得る

[例] ライドシェアビジネス



個別最適化ではなく全体最適化が可能になることによってシステムが再編され得る

[例] 省電力・需要充足・保守費用を同時最適化する社会インフラオペレーション

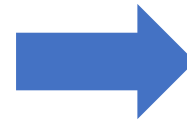


# 問題解決プロセスの進化がもたらす研究変革

## データに基づく問題解決プロセスの進化によってもたらされる効果

- 実世界で様々な起きる現象や活動の精緻かつリアルタイムの把握・予測
- 膨大な選択肢の網羅的な検証
- 大規模複雑タスクの自動実行等

これは科学研究においても極めて有効



## データ駆動型科学

様々な科学分野で活用進展

## 創薬やマテリアルズインフォマティクス

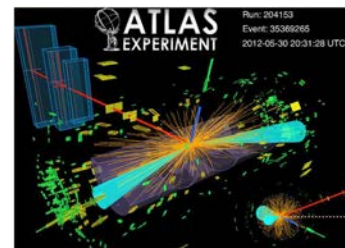
新しい薬や物性をもたらすかもしれない膨大な条件の組み合わせの中から有望なものを見つけ出すために活用されている  
(例えばCOVID-19に対するドラッグリポジショニング)

※CRDSストリーミングセミナーシリーズでは「物質・材料DX～マテリアルズインフォマティクスとその先～」「創薬DX～新薬開発のデジタル化～」にて関連動向の紹介があります

## 素粒子実験

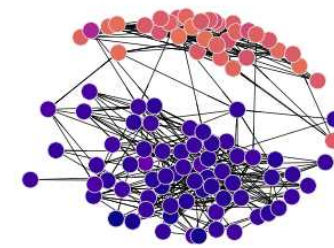
膨大な観測データ中から注目現象を選別するために活用されている  
(例えば Higgs Boson Machine Learning Challenge)

<https://www.kaggle.com/c/higgs-boson>



## 計算社会科学

社会科学においても人間行動や社会現象を観測したビッグデータの分析が進められている  
(例えば SNS上の現象分析)



[https://www.jstage.jst.go.jp/article/pjsai/JSAI2017/0/JSAI2017\\_4N1OS01a1/\\_pdf/-char/ja](https://www.jstage.jst.go.jp/article/pjsai/JSAI2017/0/JSAI2017_4N1OS01a1/_pdf/-char/ja)

# 重要な着眼点: 人間による科学的発見プロセスの限界

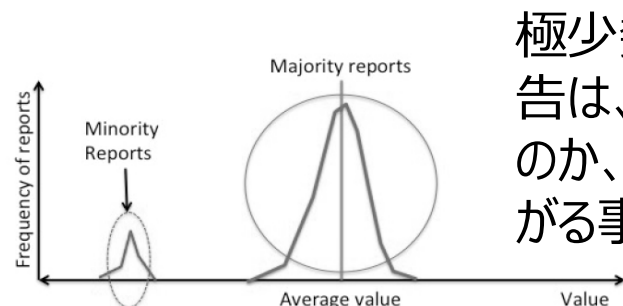
- 人間には認知能力の限界や様々な認知バイアスが存在する
- これによって科学的発見の可能性が制限されているのかもしれない

## 情報の地平線問題

論文数が幾何級数的に増加し、自分の研究に関連する全論文を読むことは困難  
人間が論文やデータを解釈する場合、必ずアクセスできない大量の情報が存在する

生命科学分野で出版される論文数は年間100万本以上

## マイノリティーレポート問題



極少数の他と異なる報告は、単なる間違いなのか、新たな発見につながる事象なのか

[参考] 北野宏明(2016)「人工知能がノーベル賞を獲る日、そして人類の未来: 究極のグランドチャレンジがもたらすもの」(人工知能31巻2号)

## 科学における様々な認知バイアス問題

研究者の期待・思い込みによる可能性の排除や見落とし、都合のよいデータの選別や解釈等

**HOW SCIENTISTS FOOL THEMSELVES — AND HOW THEY CAN STOP**

*Humans are remarkably good at self-deception. But growing concern about reproducibility is driving many researchers to seek ways to fight their own worst instincts.*

**COGNITIVE FALLACIES IN RESEARCH**

 <b>HYPOTHESIS MYOPIA</b> Collecting evidence to support a hypothesis, not looking for evidence against it, and ignoring other explanations.	 <b>TEXAS SHARPSHOOTER</b> Seizing on random patterns in the data and mistaking them for interesting findings.	 <b>ASYMMETRIC ATTENTION</b> Rigorously checking unexpected results, but giving expected ones a free pass.	 <b>JUST-SO STORYTELLING</b> Finding stories after the fact to rationalize whatever the results turn out to be.
--	--	--	---

<https://www.nature.com/news/how-scientists-fool-themselves-and-how-they-can-stop-1.18517>

# AI活用による科学的発見の可能性拡大

- 人間の持つ 認知限界や認知バイアスを超えて科学的発見の可能性を拡大 するため、網羅的探索・検証や深層学習等のAI技術を活用  
→ これから大きく進展する兆し

北野宏明(2016)やその後の講演で提唱

## Nobel Turing Challenge

2050年までに生理学・医学分野でノーベル賞級の科学的発見をできるAIシステムを作る

AIによる人間を超えた科学的発見の可能性に対するグランドチャレンジ、DARPAも注目している

丸山宏(2019)にて言及された事例

がんセンターとPreferred Networksは、4000種類以上あるExRNAの血液中の発現量を同時に見ることで、がん診断の精度を飛躍的に向上させた

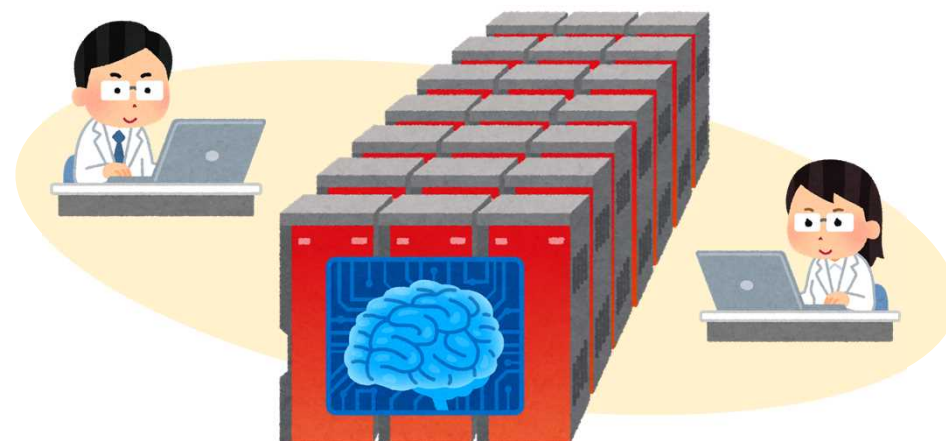
人間の頭では扱えない規模の多パラメーターの現象から規則性を発見することが深層学習には可能

- ◆ 北野宏明(2016)「人工知能がノーベル賞を獲る日、そして人類の未来：究極のグランドチャレンジがもたらすもの」(人工知能31巻2号)  
<https://ci.nii.ac.jp/naid/110010039613>
- ◆ 松尾豊(2019)「意味理会と想像」(「深層学習の先にあるもの－記号推論との融合を目指して(2)－」シンポジウムでの講演)  
[https://today.tv/contents-list/2018FY/beyond\\_deep\\_learning/01](https://today.tv/contents-list/2018FY/beyond_deep_learning/01)  
※講演の最終のパートで新しい科学の可能性に言及している
- ◆ 丸山宏(2019)「高次元科学への誘い」(CNET Japan Blog)  
[https://japan.cnet.com/blog/maruyama/2019/05/01/entry\\_30022958/](https://japan.cnet.com/blog/maruyama/2019/05/01/entry_30022958/)
- ◆ DARPA (2020)「Science 2.0」(Voices from DARPA Podcast, Episode 31)  
<https://youtu.be/FdDsNqbGpsw>



# 最後に: AI活用が研究競争力を生む時代へ

- AI・ビッグデータ活用によって問題解決プロセスが進化し、研究の世界ではデータ駆動型科学が発展、科学的発見の可能性拡大にもつながる
- AI活用が適する問題領域では、そのための技術・方法論を手にした者が、科学的発見の急拡大、研究のハイスループット化を達成し、圧倒的な研究競争力を得る可能性がある
- これに向けた技術開発・基盤整備・人材育成の重要性が高まっている



本講演のもとになった調査報告書は以下の通り

- デジタルトランスフォーメーションに伴う科学技術・イノベーションの変容 (—The Beyond Disciplines Collection—)  
<https://www.jst.go.jp/crds/report/report04/CRDS-FY2020-RR-01.html> 【4.1節】
- 研究開発の俯瞰報告書: システム・情報科学技術分野 (2019年)  
<https://www.jst.go.jp/crds/report/report02/CRDS-FY2018-FR-02.html> 【2.1.6節】

# 最後までご覧頂きありがとうございました。

## ■ DXに関するCRDS報告書はこちら



<https://www.jst.go.jp/crds/report/report04/CRDS-FY2020-RR-01.html>

## ■ CRDSの活動・成果・コラムは ウェブサイトでご覧いただけます



<https://www.jst.go.jp/crds/>



## ■ 最新情報はメールマガジン、Facebookでも配信中



<https://www.jst.go.jp/melmaga.html>



<https://fb.com/170314426446196/>



## ■ その他 お問い合わせはこちらまで



[crds@jst.go.jp](mailto:crds@jst.go.jp)