

イノベーションに求められる学び

AI ネットワークと社会システム

須藤 修 (東京大学大学院教授)

Osamu Sudoh, Prof. and PhD.

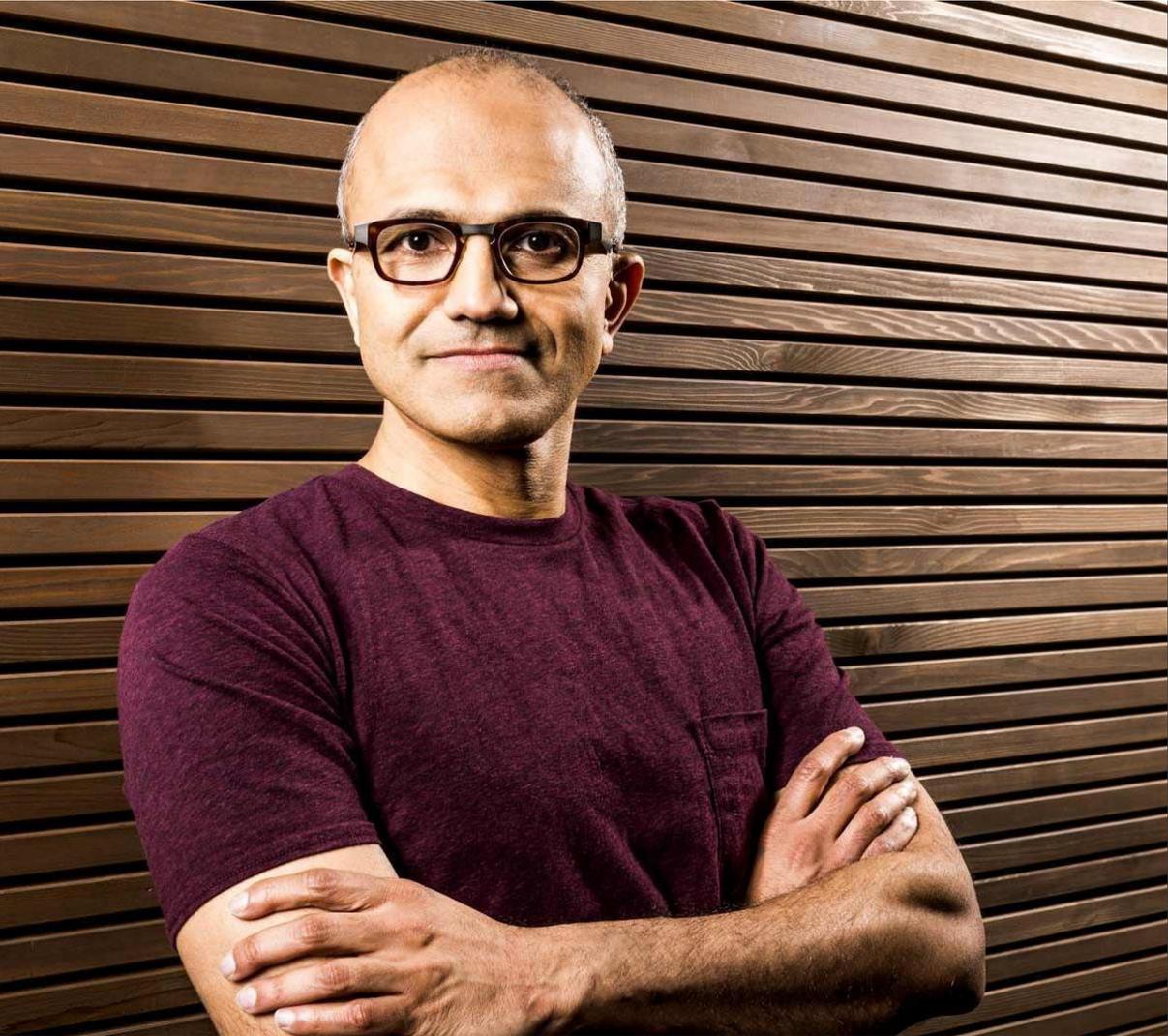
The Univ. of Tokyo

イノベーションとは何か？

■ イノベーションとは、経済活動の中で生産手段や資源、労働力などの要素をこれまでとは異なる仕方で新しく結合させて価値を生み出すことである。

✓ Joseph A. Schumpeter, *The Theory of Economic Development*, 1926





I believe

over the next decade...
intelligence will become
ambient... made possible by
an ever-growing network of
connected devices, incredible
computing capacity from the
cloud, insights from big data,
and intelligence from
machine learning.

Satya Nadella
CEO, Microsoft

提供 : Microsoft, 2016

Satya Nadella, MS CEO, states on AI, 2016

- マイクロソフトAI開発原則を発表
 - 人間の代替よりも人間の能力拡張。
 - 雇用対策は今から議論しなければならない。
 - 大きな変化はすぐ近くに来ている。
-
- 総務省AIネットワーク化検討会議中間報告(座長:須藤)でAI開発原則案発表、2016年4月
 - 高市総務大臣AI開発原則発表、G7情報通信大臣会合(2016年4月、高松市)
 - Partnership on AI, Tenetsを公表(2016年9月28日)

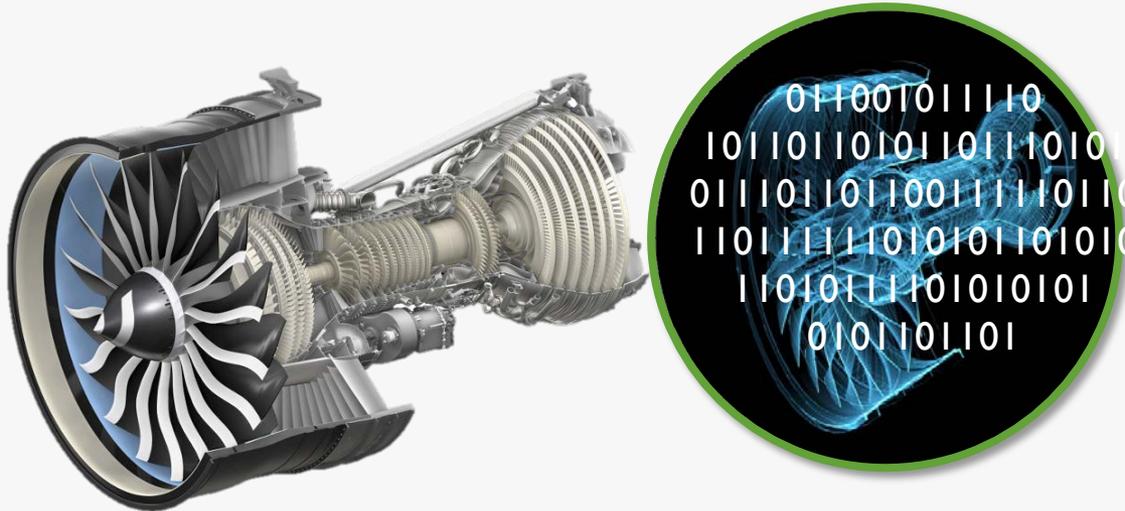
Predix Solves the Hardest IoT Challenges

How to connect securely to any asset?

How to manage and analyze massive data sets cost-effectively?

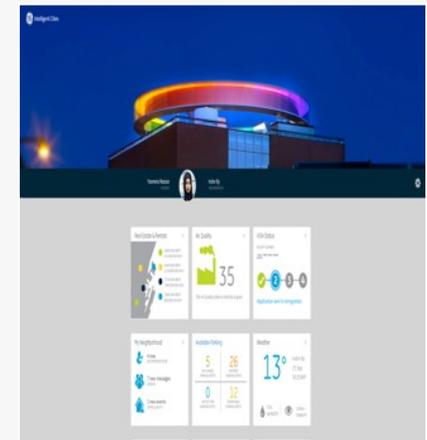
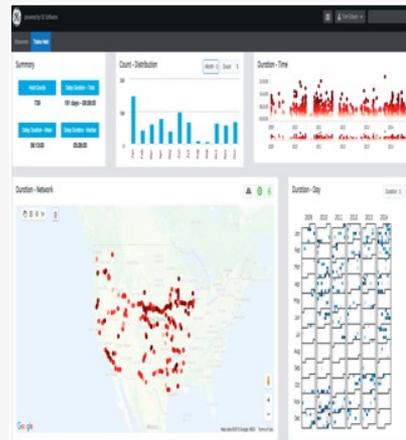
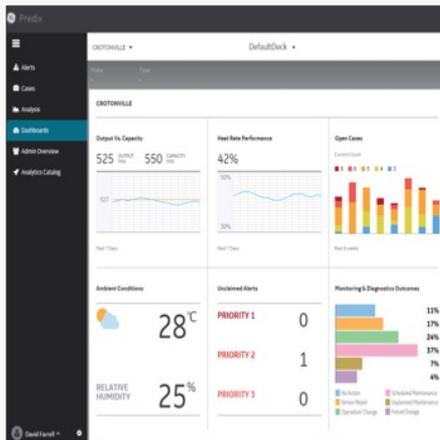
How to leverage modern app dev approaches to innovate faster?

How to protect our data and comply with regulations?



How do we integrate our operations technologies and teams with our IT platforms and teams?

GE's Digital Industrial Journey



OIL & GAS

- Maximize Production
- Predictive Maintenance
- Remote Collaboration
- Reduced Risk
- Environmental Control

POWER GENERATION

- Maximize Production
- Longer Repair Intervals
- Reduce Emissions
- Predictive Maintenance
- Longer Asset Life

POWER DISTRIBUTION

- Revenue Protection
- Meter Health
- Power Quality
- Load Forecasting
- Predictive Maintenance

WIND

- Maximize Farm Power
- Wind Wake Protection
- Outage Detection
- Continuous Operation

WATER

- Operational Integrity
- Minimize Water Use
- Control Emissions
- Minimize Cost

AVIATION

- Maximize Fuel Use
- Risk Management
- Predictive Maintenance
- Efficient Operations
- Customer Satisfaction

RAIL

- Maximize Fuel Use
- Enhanced Operation
- Network Velocity
- Predictive Maintenance
- Supplier Collaboration

HEALTHCARE

- Patient Experience
- Improved Hand Hygiene
- Cost Reduction
- Efficient Operations
- Regulatory Compliance

MANUFACTURING

- Cost Reduction
- Consumer Protection
- Efficient Operations
- Regulatory Compliance
- Predictive Maintenance

MINING

- Maximize Production
- Efficient Operations
- Safe Operations
- Predictive Maintenance

Connected Mobility Lab: Industry Research Project building the base for future Mobility Services (MaaS)

SIEMENS



Publicly funded research project

Centrum Digitalization.Bavaria (5 M€)
Lighthouse „Connected Mobility“
Living Lab
Local lab in Munich:
Platform / Data / Connected Vehicle Fleet /
Services / Apps



Verticals Partnering

 Automotive / Mobility
 Infrastructure / Traffic
 Location / Map / Analytics



Integrate Research

- Associated project with TUM
- Joint use of Connected Mobility Lab
- Transfer at dedicated synchro points
- Additional public funding (5 M€)

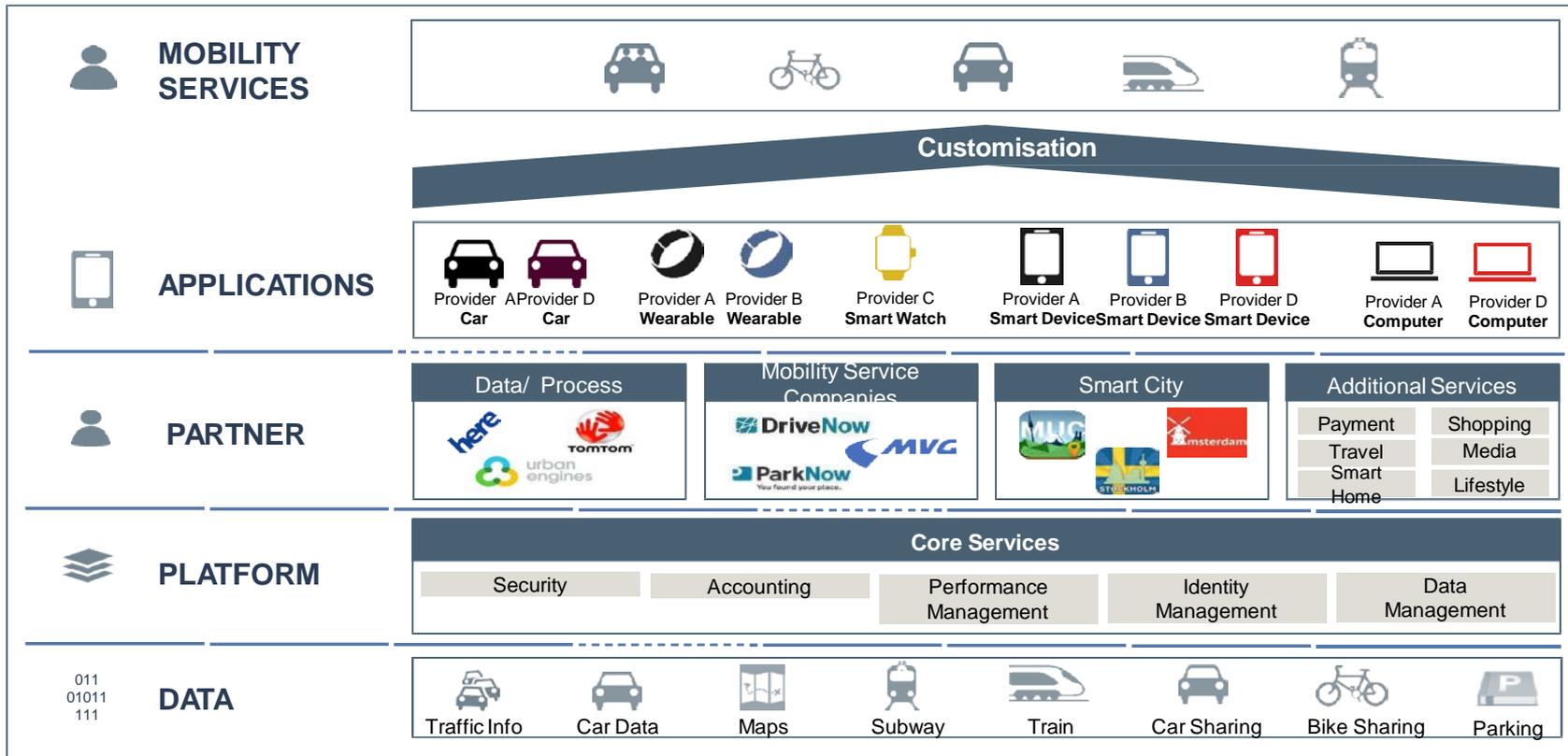


Extended Partner Network

 Startup-Scene
 Cities / Public Transp.
 Future Mob. Research Lab

CML – an open Ecosystem, which offers the framework to connects all relevant partners (e.g. cities, mobility providers, car manufactures, etc.)

SIEMENS



提供: SIEMANS, 2016



Watson on Bluemix



IBM® Watson™ Services for Bluemix™

Rapidly prototype and build cognitive apps in the cloud



出所: IBM, 2016

IBMのコグニティブ・コンピューティング

知識を活用しビジネス課題を解決

Watson ソリューション フレームワーク

活用シーン

照会応答
(Engagement)

探索・発見
(Discovery)

意思決定支援
(Decision)

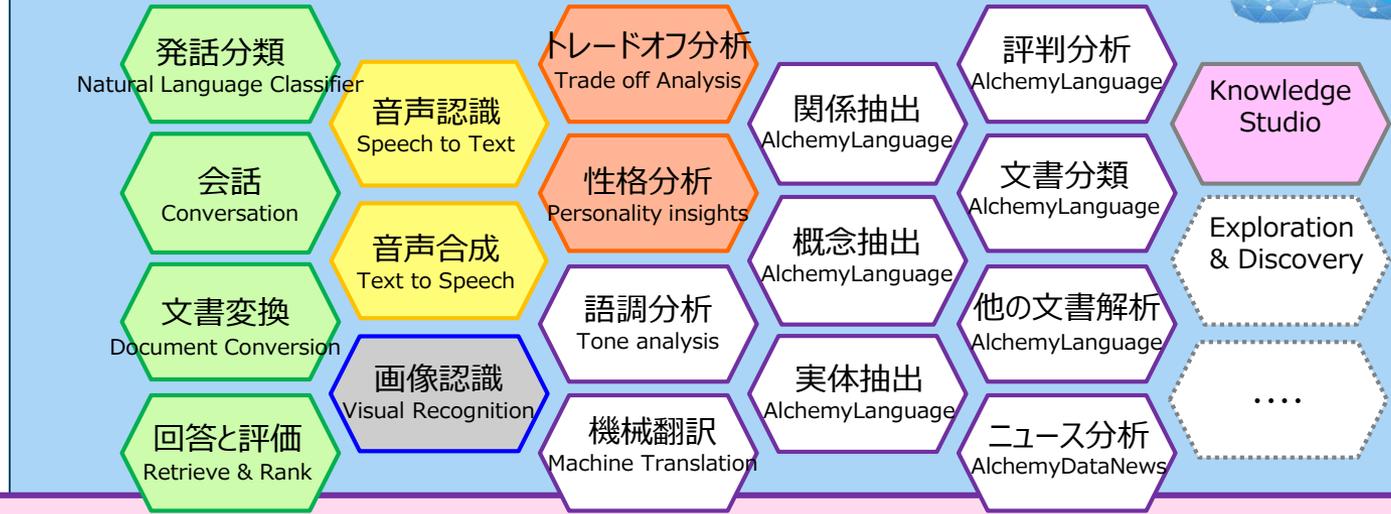


考える
読む
見る
聴く
話す

Watsonの
提供する機能



知識ベース



Watson Explorer 自然言語処理、知識表現・知識ベースの構築・分析・探索機能、機械学習



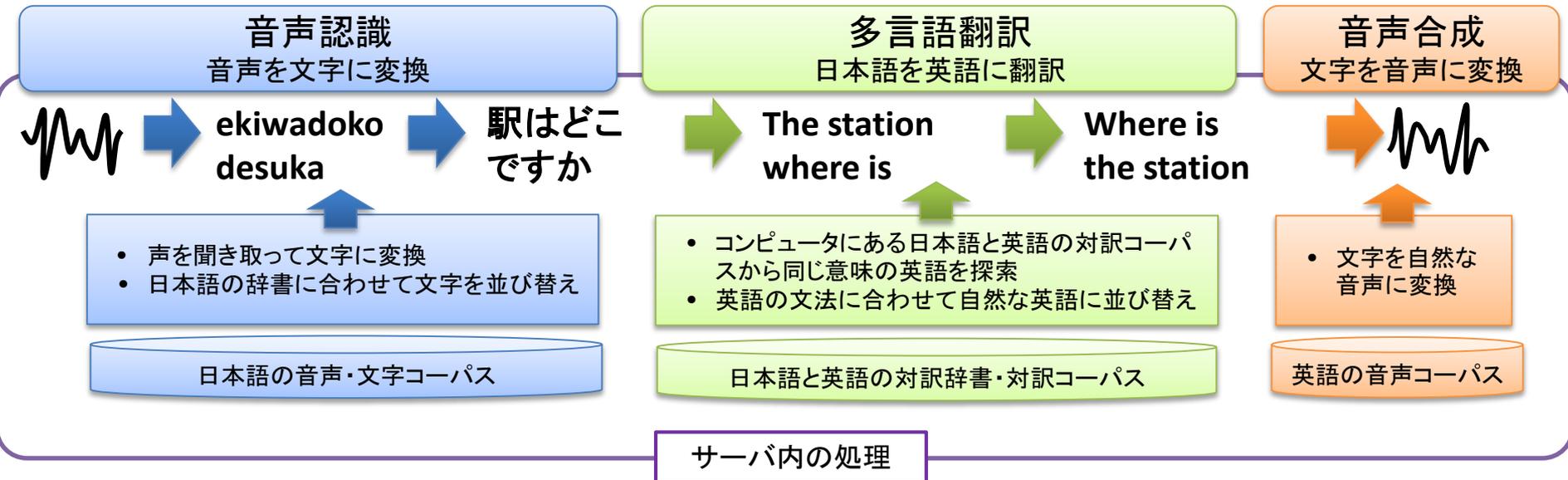
※ Watson Developer Cloudのサイトから、最新の追加サービスと更新をご確認ください。

<https://www.ibm.com/smarterplanet/us/en/ibmwatson/developercloud>

コーパスとディープラーニングを使った
多言語音声翻訳

VoiceTra

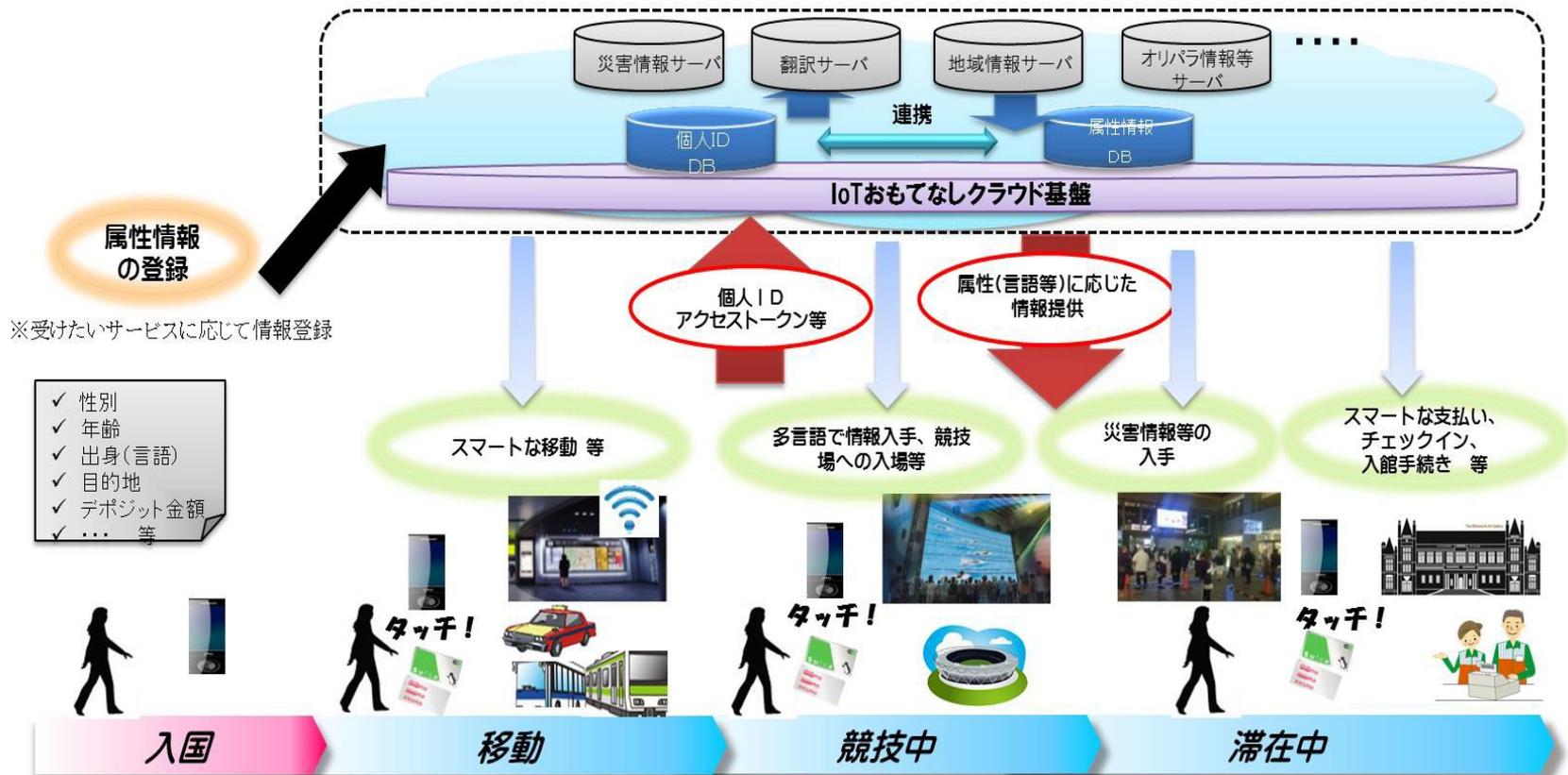
多言語音声翻訳の仕組み



コーパス: 自然言語の文章を品詞など文の構造の注釈をつけて構造化したものを大規模に集積したもの

観光 <おもてなしクラウド>

- スマートフォンや交通ICカードを活用した、個人の属性(母国語や障がいの程度等)に応じた情報提供を実現するため必要となる共通クラウド基盤。訪日外国人の方が、入国時から滞在・宿泊、観光、出国まで、ストレス無く快適に過ごすことが可能となる環境の実現を目指す。



□ 第2次人工知能ブーム

- ✓ エキスパートシステム
- ✓ Ontology, Metadata

□ 第3次人工知能ブーム(現在)

- ✓ Machine Learning Algorithms
- ✓ Cloud Computing
- ✓ 特徴量(ものごとを適切に表現するための変数)
- ✓ Deep Learning(ex.画像認識、自然言語処理), Python

□ ビッグ・プレイヤー

- IBM Watson
- Google TensorFlow, Cloud Machine Learning
- Microsoft Azure Machine Learning

Please cite this paper as:

OECD (2016), "Research Ethics and New Forms of Data for Social and Economic Research", *OECD Science, Technology and Industry Policy Papers*, No. 34, OECD Publishing, Paris.
<http://dx.doi.org/10.1787/5jln7vnpxs32-en>



OECD Science, Technology and Industry
Policy Papers No. 34

Research Ethics and New Forms of Data for Social and Economic Research

OECD

16か国から24名の研究者が招集され、検討が行われた(2014—2016年)。
日本からは佐藤一郎教授(NII)と須藤修(Univ. of Tokyo)が参画した。

報告書2016

A I ネットワーク化の影響とリスク

— 智連社会（WINS(ウインズ)）の実現に向けた課題 —

平成28年6月20日

A I ネットワーク化検討会議

AIネットワーク化検討会議

1 目的

2040年代を見据え、AIネットワーク化に向け、目指すべき社会像及びその基本理念を検討するとともに、インパクトスタディ及びリスクスタディを行い、当面の課題及び今後注視し又は検討すべき事項を整理することを目的とする。

2 検討体制

【座長】 須藤 修 東京大学大学院情報学環教授

【顧問】 村井 純 慶應義塾大学環境情報学部長

【構成員】 理工学系及び人文・社会科学系の有識者 計37人（座長及び顧問を含む。）

3 スケジュール

2月 2日 第1回会合
4月15日 中間報告書

（ 親会 5回開催
分科会（経済、社会・人間、法・リスク）等 13回開催 ）

目指すべき社会像及び基本理念、インパクトスタディ・リスクスタディ、当面の課題

6月20日 報告書2016

「智連社会」における人間像、評価指標、リスク・シナリオ分析、今後の課題

【参考】4月29日・30日 G7香川・高松情報通信大臣会合

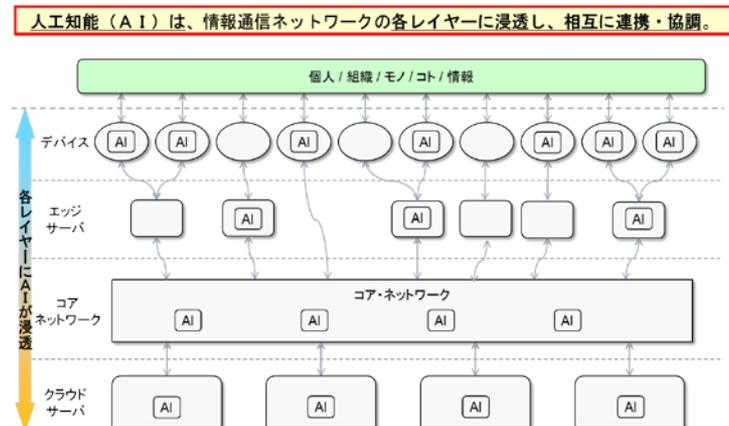
「AIネットワーク化」の進展段階

① AIが、他のAIとは連携せずに、インターネットを介するなどして単独で機能し、人間を支援

② AI相互間のネットワークが形成され社会の各分野における自動調整・自動調和が進展

- ネットワーク上に用途の異なる多様なAIが出現。
- 複数のAIを取りまとめる能力を有するAIも出現。
- 複数のAIが相互に連携・協調。

(例) ・産業機械と部材の連携、サービスロボットとセンサの連携
・交通、物流、オフィス業務、生活環境等の自動調整



③ 人間の潜在的な能力がAIネットワークシステムにより引き出され、身体的にも頭脳的にも発展

センサ、アクチュエータ、人間及び人工知能が連携

感覚器官の能力向上

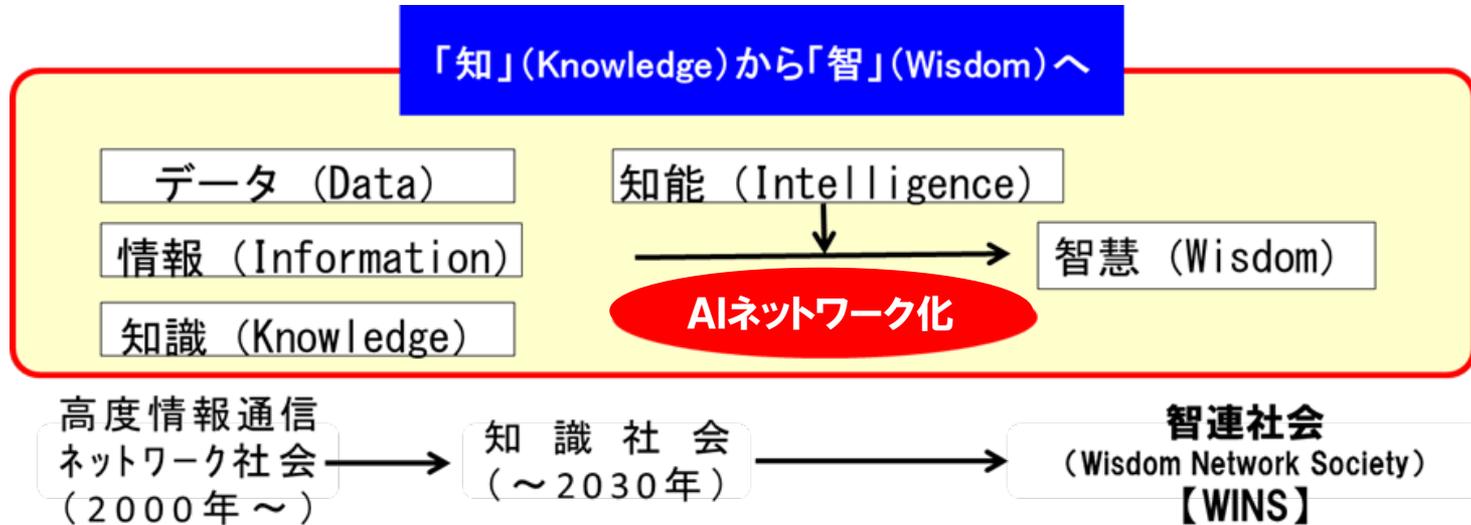
身体機能の能力向上



- (例) ・脳情報を外部に出力
・ヒトの意思により、義手・義足やロボットを操作
・遠隔地の出来事を仮想体験
→ 遠隔地のロボットを操作して協働

④ 人間とAIネットワークシステムが共存

「知」から「智」へ



AIネットワーク化の進展段階

- ① **AIが**、他のAIとは連携せずに、インターネットを介するなどして**単独で機能し**、人間を支援
- ② **AI相互間のネットワークが形成され**、社会の各分野における自動調整・自動調和が進展
- ③ **人間の潜在的能力がAIネットワークシステムにより引き出され**、身体的にも頭腦的にも発展
- ④ **人間とAIネットワークシステムとが共存**

「データ」・「情報」・「知識」・「知能」・「智慧」の関係

データ (Data)	断片的な事実、数値、文字
情報 (Information)	データの組み合わせに意味を付与したもの
知識 (Knowledge)	データ・情報の体系的集積
知能 (Intelligence)	データ・情報・知識を解析することにより、新たなデータ・情報・知識を創造する機能
智慧 (Wisdom)	データ・情報・知識に基づき、知能を活用することにより、物事に対処する能力

AIネットワーク化の進展が産業構造にもたらす影響①

異業種間の融合の進展

A I ネットワーク化の進展に伴い、

- ① 隣接業界への進出による新たな製品・サービスやビジネスモデルの創出
 - ② 異業種間の連携による新たな製品・サービスやビジネスモデルの創出
- といった形での異業種間の融合の進展が見込まれる。

これらの事業はA I ネットワークシステムによりデータを利活用することによって成立するもの。

産業のあらゆる分野において情報通信産業化が進行していくと評価することも可能。

これらの変化を牽引できるのはベンチャー企業等新しく小さい企業。

- ベンチャー企業等小さい企業の起業は、そのセットアップコストの大幅な低下等により、今後一層容易になっていく。
- 産業連携や新たなビジネスモデルの企画立案に対応できる人材が工夫すれば、これらの事業といえども、世界を相手にして開始可能。

AIネットワーク化の進展が産業構造にもたらす影響②

異業種間の融合の進展の例

<隣接業界への進出による新たな製品・サービスやビジネスモデルの創出の例>

- ① 製造業者が、自らの製品から稼働状況等に関するデータを収集し、収集したデータに基づいてA I ネットワークシステムを用いて製品の状態等を常時分析し、製品の保守等アフターサービスに活かすことにより、サービス業に進出するケース
- ② 農業等第一次産業に属する企業が自らの産品への需要及び供給に関するデータをA I ネットワークシステム上で活用することが可能となることにより、自ら流通の業務に進出するケース

<異業種間の連携による新たな製品・サービスやビジネスモデルの創出の例>

- ① 自動車製造業者と保険業者とが連携して、ドライブレコーダと接続しているA I ネットワークシステムを活用することにより、走行履歴のデータの分析結果に基づいて自動車保険料の割引を実施するケース
- ② 豊富なデータを保有する大企業とデータ分析ベンチャーが連携して、SNSの情報及びPOS（購買）データの双方をA I ネットワークシステムで分析することにより、新たな製品開発やマーケティング手法を開発するケース

AIネットワーク化の進展が雇用にもたらす影響①

AI やロボット等による技術的代替

AI ネットワーク化の経済効果は極めて大きく、新たな雇用も創出されるものと考えられているが、同時に、定型的業務のみならず、知的業務といえども、AI ネットワークシステムによって代替することが技術的には可能。



人間に残される仕事の特徴としてはクリエイティビティ、マネージメント及びホスピタリティが指摘されており、人間はアイデアを生み出す仕事や人間相互間の高度なコミュニケーション能力（リーダーシップ等）を必要とする仕事を担う。



- AI に代替されない能力を身に付けさせるよう、教育の在り方の抜本的な検討が必要。
- 現状では、人間の仕事を奪うのは、AI を駆使する人間である。AI を使いこなすためのプログラミング（言語）能力もさることながら、今後は再編成能力（※）が重要。

※どこまでをコンピュータに任せ、どこまでを人間にやらせるかという判断に要する能力

AIネットワーク化の進展が雇用にもたらす影響②

就労者と企業との関係に新たな可能性

- 場所的・時間的拘束を軽減するテレワークや、言語的制約を軽減する自動翻訳は、自営的就労（※）の可能性を拡大。
※ 場所的・時間的拘束が少なく、自分で仕事のオーガナイズができる働き方
- 企業内の仕事の再編成が行われ、そのかなりの部分の外製化（特に他の自営的就労者等へのクラウド・ソーシング）が進むため、これまでの労働法制が前提としてきた、職種、勤務地、勤務時間等を限定しない無期雇用たる正社員を中心とする企業組織や雇用の在り方に、抜本的な影響が及ぶ。



正社員を中心とする企業組織や雇用の在り方を前提としてきた伝統的な労働法制については、抜本的な見直しが必要。

「智連社会」における人間像(1)

<人間像の提示>

- AIネットワーク化の進展が産業構造や雇用にもたらす様々な影響や人間の潜在的能力の拡張可能性は、社会を構成する人間の生き方を左右し得るものと考えられる。
- 「智連社会」の実現を目指して取り組む場合においては、各人の「智慧」(智)が相互に連携し合い、相互に協調し合うことによる人間の生き方の変化、さらには、人間の在り方の変化も念頭に置くことが必要と考えられる。
- 「智慧」(智)の「連携」及び「協調」は、人間がAIネットワークシステムとシームレスに連携し、人間同士もAIネットワークシステムを介してシームレスに連携することまでが可能となるという状況における「連携」及び「協調」も念頭に置くことが必要となる。

・人間がインテリジェントICTと共存し

人機共存

・データ・情報・知識を自由かつ安全に創造・流通・連結して智のネットワークを構築することにより

総智連環

・あらゆる分野におけるヒト・モノ・コト相互間の空間を越えた協調が進展し

協調遍在

もって創造的かつ活力ある発展が可能となる社会

<「智連社会」の実現が期待される将来の社会における人間を取り巻く状況(主な見解)>

- 定型的業務のみならず、知的業務といえども、AIネットワークシステムによって代替することが技術的には可能となり、就労人口が減少しているのではないか。
- 雇用の変化(正社員の減少、自営的就労者の増加、テレワークの増加等)が生じているのではないか。
- 所得格差が拡大し、生存可能賃金を下回る賃金水準となる可能性が生ずるに至っているのではないか。
- 今までにない産業や仕事などが生み出され、成長や発展、自己実現等に関し様々なチャンスが出てくる社会となる可能性があるのではないか。

「智連社会」における人間像(2)

<価値観の多様化>

- 仕事や雇用のみには重きを置く価値観だけではなく、価値観の多様化が実現すれば、智連社会における人間の生き方が多様化することにもつながるものと考えられる。
- 生活等を賄う資金を確保するための手段の一つとして、**ベーシックインカム**の導入を検討すべきとの見解が示されたが、様々なメリットとデメリットの双方が指摘されており、多面的な検討が必要である。

<仕事で求められる能力>

- A I ネットワーク化の進展に伴って変化する社会に適応していくための能力開発・技能習得が求められる。
- 求められる能力が単なる認知能力（読み、書き、そろばん）から非認知能力（事業再編成能力、コミュニケーション能力、全体をデザインする能力等）にシフトすることが示されているが、A I ネットワーク化の進展を見据えて、これらの能力の育成のために、初等教育から高等教育までを通じて体系的な教育・訓練が可能となるよう教育改革が必要となるであろう。

<人間の潜在的能力の拡張>

- A I ネットワーク化の進展段階のうち第三段階において、人間の身体とA I ネットワークシステムとが連携することにより、人間の潜在的能力が拡張し得ることが展望されている。
- 人間とサイボーグやロボットが共存する社会においては、まさに**人間とは何か、人間のあるべき姿とは何かといった人間の存在に関する根本問題の検討が迫られる可能性がある。**
- **解読された人間の脳情報がインターネット等を介してコンピュータにアップロードされたり、複製されたり、ロボット等にダウンロードされたりする可能性も視野に入れるならば、人間の生死に関する根本問題にまで立ち入った倫理的な検討までもが必要となり得る。**

第4章 リスク・シナリオの具体例 (1)機能に関するリスク(1/2)

リスクの種類	シナリオ上想定されるリスクの内容	発生時期	生起確率	被害の規模	二次的(派生的)に生ずるリスク	リスク評価	リスク管理	リスク・コミュニケーション
セキュリティに関するリスク	ロボット自身がハッキング攻撃されることにより、踏み台として利用され、情報が流出したり、ロボットが不正に操作されるリスク	進展段階1	高	中	プライバシー・個人情報に関するリスク、犯罪のリスク	ロボットのセキュリティ上の脆弱性等の評価	事後検証のためのロギングの実装、脆弱性の発見・対処	生じたインシデントに関する情報共有と対応策についての説明
セキュリティに関するリスク	ロボットに関係するクラウド等AIネットワークシステムがハッキング攻撃されることにより、情報が流出したり、ロボットが不正に操作されるリスク	進展段階2	高	大	プライバシー・個人情報に関するリスク、消費者等の権利利益に関するリスク、犯罪のリスク	情報流出による影響の評価、クラウドのセキュリティ上の脆弱性等の評価	事後検証のためのロギングの実装、脆弱性の発見・対処、必要に応じて結合テストの追試	生じたインシデントに関する情報共有と対応策についての説明
情報通信ネットワークシステムに関するリスク	ネットワークの遅延や停止によりロボットが動作しなくなったり、想定外の動作をするリスク	進展段階1	高	中	セキュリティに関するリスク、制御喪失のリスク、事故のリスク	どの部分の遅延、停止によってどのような動作が起こりうるか等の評価	原因把握、必要に応じて結合テストの追試	生じたインシデントに関する情報共有と対応策についての説明
情報通信ネットワークシステムに関するリスク	AIネットワーク化の進展により、フレキシブルなモジュール間連携が可能となる反面で、想定外のネットワークングにより、想定外の処理が行われ、ロボットが想定外の動作をするリスク	進展段階2	低	大	セキュリティに関するリスク、事故のリスク、制御喪失のリスク、事故のリスク	どの部分のネットワークングがどのような影響を及ぼすか等の評価	原因把握、必要に応じて結合テストの追試	生じたインシデントに関する情報共有と対応策についての説明

第4章 リスク・シナリオの具体例 (1)機能に関するリスク(2/2)

リスクの種類	シナリオ上想定されるリスクの内容	発生時期	生起確率	被害の規模	二次的(派生的)に生ずるリスク	リスク評価	リスク管理	リスク・コミュニケーション
不透明化のリスク	ロボットのインターフェースの不備により、動作に至る過程や根拠を確かめることが困難になるリスク	進展段階1	高	小	セキュリティに関するリスク、制御喪失のリスク、事故のリスク	通常時のみならず、異常時についてもロボットがインタフェースとしてどれほど機能するか等の評価	事後検証のためのロギングの実装	通常時からの利用者と開発者による情報の共有等
不透明化のリスク	ネットワーク上で複数のAIが多重かつ複雑に連携してロボットを操作する場合、不確実性が増大し、動作に至る過程や根拠がブラックボックス化するリスク	進展段階2	中	大	セキュリティに関するリスク、制御喪失のリスク、事故のリスク	どの部分のネットワークキングがどのような影響を及ぼすか等の評価	事後検証のためのロギングの実装、原因把握、必要に応じて結合テストの追試	通常時からの利用者と開発者による情報の共有等
制御喪失のリスク	ファームウェアの乗っ取りや不正なアップデートなどにより、ロボットが想定外の動作をし、制御が喪失するリスク	進展段階1	低	大	復旧までロボットの機能を代替することが困難となるリスク。ロボットの制御喪失が他のロボット等の制御喪失や停止に波及するリスク	実際の乗っ取りがどれほどの危険を生じるか、ファームウェアアップデートにおける脆弱性等の評価	暴走したロボットをネットワークから切断、停止(再起動)した後、復旧	生じたインシデントに関する情報共有と対応策についての説明
制御喪失のリスク	自動走行車(レベル3を想定)の運転中に機能不全が生じた場合に、運転者の技能低下や機械の不調などにより、運転者が操作に介入することができず、制御不能に陥るリスク	進展段階1(2020年代)	中	中	事故のリスク、移動手段喪失のリスク	運転者の生命身体へのリスク、関連車両群や交通システム等への影響等の評価	運転者の技能維持のための試験・講習等、リスク改善に向けた作業・モニタリング等	運転者への啓発、緊急事の情報共有等

第4章 リスク・シナリオの具体例 (2)法制度・権利利益に関するリスク(1/3)

リスクの種類	シナリオ上想定されるリスクの内容	発生時期	生起確率	被害の規模	二次的(派生的)に生ずるリスク	リスク評価	リスク管理	リスク・コミュニケーション
事故のリスク	レベル3の自動走行車の運転時に運転者がハンドルから手を離して乗ることにより、緊急時の対応が困難になるリスク	進展段階1	中	中	制御喪失のリスク	運転者の生命身体へのリスク、関連車両群や交通システム等への影響等の評価	運転者の技能維持のための試験・講習等、リスク改善に向けた作業・モニタリング等	運転者への啓発、緊急事の情報共有等
事故のリスク	自動走行車が、ネットワークを通じて、誤った情報を共有したり、共鳴することで交通システムが麻痺することにより、事故が生じるリスク	進展段階2	低	大	情報通信ネットワークシステムに関するリスク、セキュリティに関するリスク	ネットワークの機能不全による交通システムの麻痺のリスク等の評価	システムダウンした場合の手動による対応の準備等	システムダウン時の対応の周知、システム状況の把握
犯罪のリスク	親しみのある見た目の人型ロボットが、オレオレ詐欺の「受け子」や「出し子」など人間の代替物として犯罪に悪用されるリスク	進展段階1	中	中	消費者等の権利利益に関するリスク	ロボットに対する人間の信頼感等の評価	ロボットの登録制やIDの表示等	ロボットを利用した犯罪手法に関する情報共有
犯罪のリスク	個人Aの脳と連携したAI・ロボットが個人Bにより不正に操作され、個人Bが個人Aを利用して犯罪を実行するリスク	進展段階3	不確実	中	人間の尊厳と個人の自律に関するリスク、民主主義と統治機構に関するリスク	人間の脳の判断・身体動作が外部から不正に操作される脆弱性等の評価	脳と連携したAIのセキュリティの強化、ロギング(外部との通信等の記録)、責任の帰属の在り方の検討等	利用者に対する説明、相談窓口・通報制度の整備

第4章 リスク・シナリオの具体例 (2)法制度・権利利益に関するリスク(2/3)

リスクの種類	シナリオ上想定されるリスクの内容	発生時期	生起確率	被害の規模	二次的(派生的)に生ずるリスク	リスク評価	リスク管理	リスク・コミュニケーション
消費者等の権利利益に関するリスク	愛玩用の犬型ロボットの飼い主のリテラシー不足などにより、ロボットのアップデートが確実になされなかったため、ロボットが遠隔操作ウィルスに感染して、悪用され、空き巣に入られたり、情報が漏洩するなどの被害が生ずるリスク	進展段階1	高	中	セキュリティに関するリスク、プライバシー・個人情報に関するリスク、犯罪のリスク	セキュリティ機能が更新されないことにより安全な利用が困難となる蓋然性、被害の規模等の評価	自動アップデートの整備、ウィルス感染時の停止・ネットワークからの切断等	セキュリティ機能等の更新についての状況把握・情報共有
消費者等の権利利益に関するリスク	愛玩用の犬型ロボットが歌うサービスを提供していた会社が倒産したため、サービスが継続できず、ロボットが歌わなくなり、ショックを受けた飼い主の高齢者の健康が悪化するリスク	進展段階2	中	中		AIネットワークシステムの連携により導入される他者サービス又は連携により生じたサービスの継続性等の評価	データ・ポータビリティの確保等	末端利用者の連携状況の把握、AIネットワークシステムの状況について把握するための定型化等(連携状況をブラックボックスにしない)
プライバシー・個人情報に関するリスク	サービス・ロボットのプロファイリングにより健康状態等に関する(差別に繋がる、誤った)情報が伝播するリスク	進展段階1	不確実	不確実		プロファイリングによる差別に繋がる情報の伝播のリスク等の評価	人間が介在しないプロファイリングの監視又はプロファイリング結果の修正手段の確保	プロファイリング結果等へのアクセスの確保
プライバシー・個人情報に関するリスク	サービス・ロボットとドローンがネットワークを通じて連携し、利用者とロボットとの会話に関係する商品をドローンが自動的に配送するサービスにより、望まない商品が配送されるが、適切な修正が不可能であるリスク	進展段階2	中	中	消費者等の権利利益に関するリスク	AIネットワークシステムの連携により導入される他者サービス又は連携により生じたサービスの透明性等の評価	苦情窓口の開設、自宅へのロボット・ドローン等のアクセス制限等	透明性に関するルール策定及びモニタリング、透明性確保のための連絡先フォーマットの統一化等

第4章 リスク・シナリオの具体例 (2)法制度・権利利益に関するリスク(3/3)

リスクの種類	シナリオ上想定されるリスクの内容	発生時期	生起確率	被害の規模	二次的(派生的)に生ずるリスク	リスク評価	リスク管理	リスク・コミュニケーション
人間の尊厳と個人の自律に関するリスク	ロボットにより撮取する情報等を操作されることにより、利用者の意思決定や判断のプロセスが操作されるリスク	進展段階2	中	大	民主主義と統治機構に関するリスク	利用者の意思決定や判断に介在する蓋然性・程度等の評価	システム設計段階における指標の確立	AI・ロボットによる個人の意思決定や判断への影響の注視・啓発、リテラシーの涵養等
人間の尊厳と個人の自律に関するリスク	遺伝子等を元に亡くなった人を再現するロボットが人間の尊厳との関係で問題となるリスク	進展段階3	低	不確実(価値判断を伴う)		肉体以外は人のクローンに限りなく近い人工物の作成に関する意識調査等	人のクローンに近づく人工物につき原則禁止とするか原則自由として事後的規制を課すかについての事前の議論	倫理的問題についてのステークホルダー間での議論
民主主義と統治機構に関するリスク	トレイグジスタンス・ロボットにより外国人が入国審査を受けることなく「上陸」することが可能となり、出入国管理制度が機能不全に陥り、テロリスト等が流入するリスク	進展段階1	低	不確実	犯罪のリスク	国外からのトレイグジスタンス・ロボット操作による影響等の評価	アクセスログの記録、緊急時のアクセスの制限等	国際的な連携・情報共有等
民主主義と統治機構に関するリスク	人間に投棄された「野良ロボット」が徒党を組んで人間に対して参政権等の権利付与を要求するリスク	進展段階4	低	大	制御喪失のリスク	権利付与を求めるほど高度な知能を持ちえるのか、持ち得たとして実際に権利付与を求めるか等の評価	「野良ロボット」が生じないよう登録制等の検討、人間に反乱するおそれのある人工知能の開発の事前の制限等	ロボットと人間の関係についてのステークホルダー間での議論

今後の課題

研究開発の原則・指針の策定

研究開発の原則及びその説明から構成される指針（A I 開発ガイドライン（仮称））を国際的に参照される枠組みとして策定することに向けての検討が必要。

【原則の策定・解釈に当たっての基本的な考え方】次に掲げる考え方を基本的な考え方として前文等に記述すべき。

- ①
 - ・ A I ネットワークシステムの恵沢が万人により享受され、
 - ・ 人間の尊厳と個人の自律が保障され、
 - ・ A I ネットワークシステムの制御可能性と透明性が確保され、
 - ・ A I ネットワークシステムが安全に安心して利活用される
- ② 研究開発の進展段階に応じて、想定される各種のリスクに適時適切に対処
- ③ イノベティブな研究開発と公正な競争にも配慮しつつ、多様なステークホルダーの参画を得て、関係する価値・利益のバランスを調整
- ④ A I ネットワーク化の進展及びリスクの顕在化に応じて、適宜見直し

社会を実現

【原則の内容】次に掲げる事項を基本に検討すべき。

- ① 透明性の原則 AIネットワークシステムの動作の説明可能性及び検証可能性を確保
- ② 利用者支援の原則 AIネットワークシステムが利用者を支援、利用者に選択の機会を適切に提供するよう配慮
- ③ 制御可能性の原則 人間によるAIネットワークシステムの制御可能性を確保
- ④ セキュリティ確保の原則 AIネットワークシステムの頑健性及び信頼性を確保
- ⑤ 安全保護の原則 AIネットワークシステムが利用者及び第三者の生命・身体の安全に危害を及ぼさないよう配慮
- ⑥ プライバシー保護の原則 AIネットワークシステムが利用者・第三者のプライバシーを侵害しないよう配慮
- ⑦ 倫理の原則 AIネットワークシステムの研究開発において、人間の尊厳と個人の自律を尊重
- ⑧ アカウントビリティの原則 AIネットワークシステムの研究開発者が利用者に対するアカウントビリティを遂行

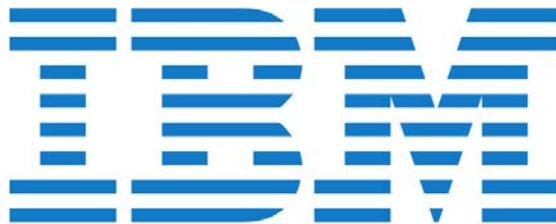
→ ガイドライン案の作成に向けて8原則のブレイクダウン
国際社会に向けてOECD等における継続的議論の必要

amazon.com



Partnership on AI
to benefit people and society

facebook



Partnership on AI (いわゆる「五社連合」)について

正式名称: Partnership on Artificial Intelligence to Benefit People and Society

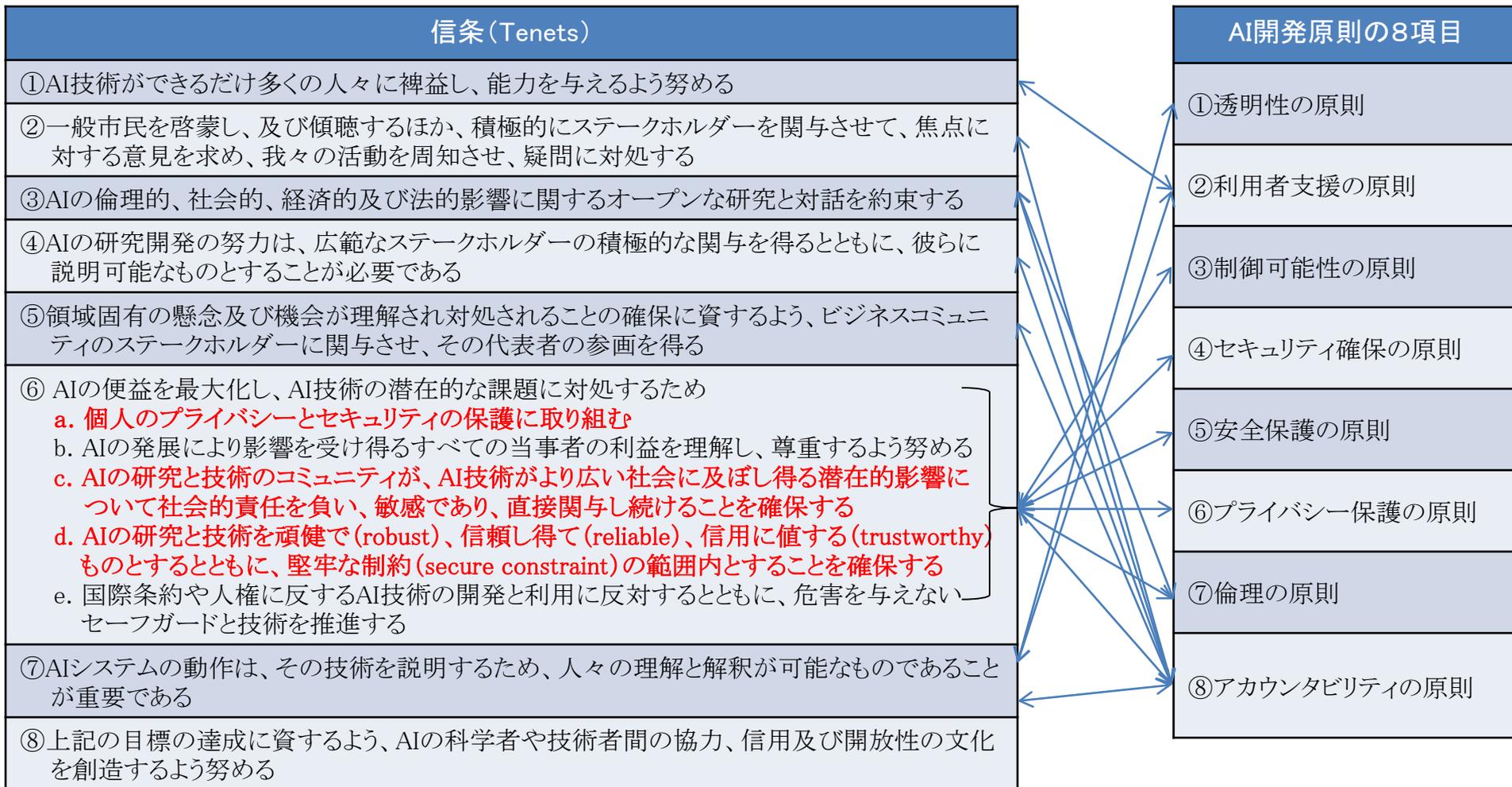
設立: 2016年9月28日(水)

構成企業: Amazon、DeepMind/Google、Facebook、IBM、Microsoft



暫定共同議長: エリック・ホロヴィッツ(Microsoft技術フェロー)及びムスタファ・スレイマン(DeepMind共同創業者)

【五社連合の設立時に公表した「信条」(Tenets)と、AIネットワーク化検討会議が「AI開発原則」につき掲げる8項目との対応関係(未定稿)】



総務省「AIネットワーク社会推進会議」

(第1回会合 2016年10月31日)

- 議長：須藤 修
- 顧問：
 - 安西祐一郎(慶応大学元塾長)
 - 長尾 真(京都大学元総長)
 - 西尾章治郎(大阪大学総長)
 - 濱田純一(東京大学元総長)
- 「AIネットワーク化検討会」を発展的に改組。
- <開発原則>と<影響評価>に焦点を当てて検討を行う。

Dr. Rita Colwell (全米科学財団元長官)

- Across the science and engineering enterprise, boundaries are increasingly difficult to distinguish between and among disciplines, especially information technology, nanotechnology, and many areas embracing biocomplexity, the complexity of life itself.
- But it is true also for the social sciences in this new era of “Big Data”, when computational capacity reaches beyond imagination.



- **The most exciting areas are in these fuzzy connections between disciplines where knowledge in one field answers questions in another field.**