

実務者が考える AI・データ活用における倫理フレームワーク



AI・データ活用のためのコンプライアンス研究会

日本データマネジメント・コンソーシアム
Japan Data Management Consortium [JDMC]

■ テーマ

- デジタル経済の加速による、かつてないデータ量・活用幅の拡大に伴う課題や今後のビジネス拡大のための着眼点について、コンプライアンスの観点からディスカッション

■ コンセプト

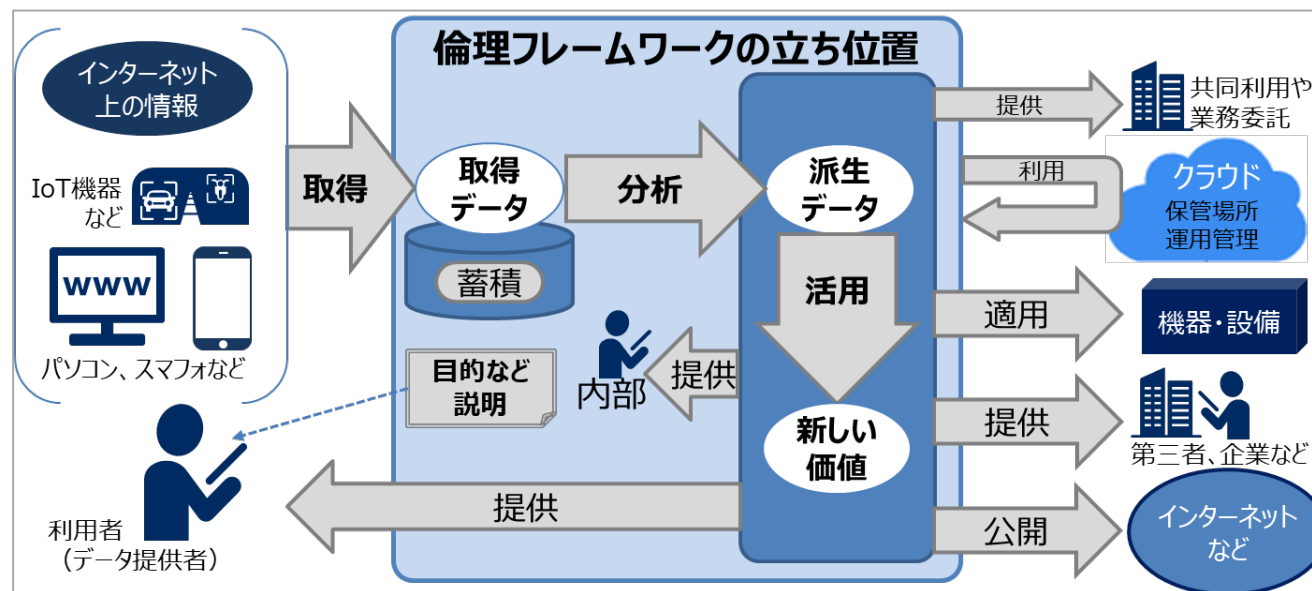
- データを戦略的に収集し適切に管理することが、データ活用の本格化に繋がる
- 誤ったデータの取り扱いをすることは経営における大きなリスクとなる
- AI・データ活用を促進していくために、セキュリティ、コンプライアンス、個人情報保護を、社内規定や社外との契約で、厳格に取り扱っていくことが必要
- 契約、コンプライアンス、個人情報保護等について、実務対応を検討とデータ活用を企業として戦略的に活用する基礎となる研究を実施

■ これまでの活動

- 2019年 当研究会発足 重要ポイントの全体像を整理
- 2020年 倫理フレームワーク、データ活用におけるつまずきポイントを検討
- 2021年 倫理フレームワークのつまずきポイントにおける対応策を検討

■ 検討にあたっての立ち位置

- データの入手から活用においては、関連する人/組織は複数あるが、倫理フレームワークは、データ提供者から取得したデータを分析し活用する人/組織向けのものである。
- 活用する人/組織が作り出した派生データを、外部へ提供したり公開することも想定している。



■ つまずきポイントとは

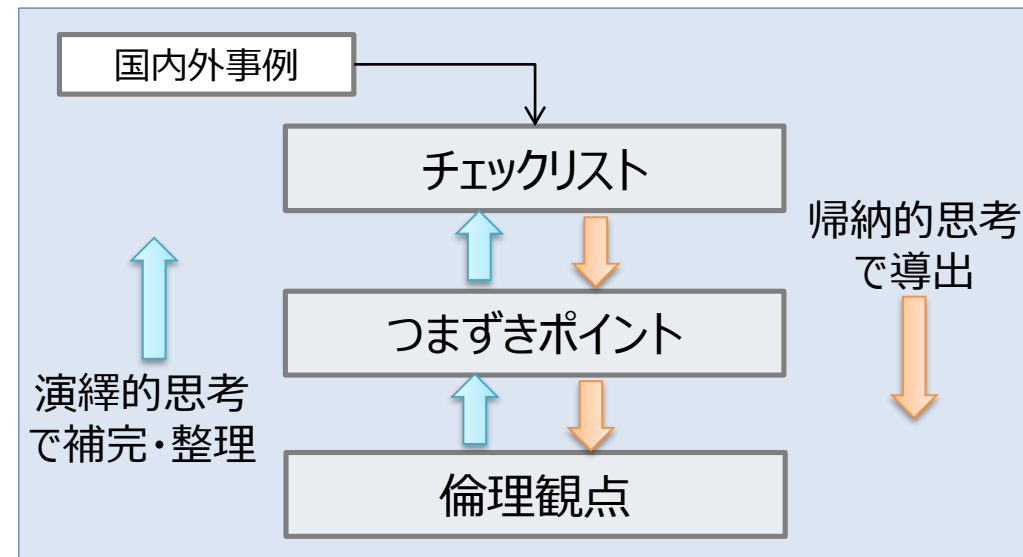
法には抵触しないが問題となり得る要因

■ 目指したこと

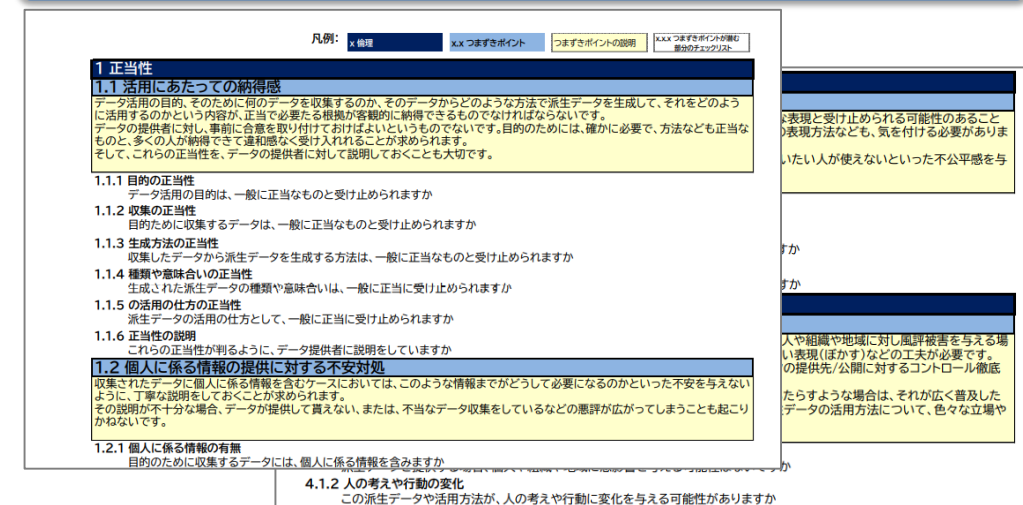
- ✓ チェックリスト形式とし、データ活用に携わる実務者自身が活用できるように
- ✓ 過去の事例だけではなく、まだ指摘されていないようなつまずきポイントについても抽出できるように

■ 倫理フレームワークの作成のアプローチ

- ① 事例を予防する「チェックリスト」項目を設定
 - ② 複数の「チェックリスト」項目に共通の要因を抽出し、「つまずきポイント」を導出
 - ③ 「つまずきポイント」の背後にある「倫理観点」を導出
 - ④ 実経験に基づく議論より、「つまずきポイント」および「チェックリスト」項目を補完
- 以上を繰り返し、ブラッシュアップのうえ、「倫理フレームワーク」としてまとめた



倫理フレームワーク



凡例: x 倫理 xx つまずきポイント つまずきポイントの説明 xxx つまずきポイントの補完・整理のチェックリスト

1 正当性

1.1 活用にあたっての納得感

データ活用の目的、そのために何のデータを収集するのか、そのデータからどのような方法で派生データを生成して、それをどのように活用するかという内容が、正当に必要な根拠が客観的に納得できるものでなければなりません。データの提供者に対し、事前に合意を取り付けておけばよいというものでないです。目的のためには、確かに必要で、方法なども正当なもの、多くの人が納得できて違和感なく受け入れられることが求められます。そして、これらの正当性を、データの提供者に対して説明しておくことも大切です。

1.1.1 目的の正当性
データ活用の目的は、一般に正当なものと受け止められますか

1.1.2 収集の正当性
目的のために収集するデータは、一般に正当なものと受け止められますか

1.1.3 生成方法の正当性
収集したデータから派生データを生成する方法は、一般に正当なものと受け止められますか

1.1.4 種類や意味合いの正当性
生成された派生データの種類や意味合いは、一般に正当に受け止められますか

1.1.5 の活用の仕方正当性
派生データの活用の仕方として、一般に正当に受け止められますか

1.1.6 正当性の説明
これらの正当性が判るように、データ提供者に説明をしていますか

1.2 個人に係る情報の提供に対する不安対処

収集されたデータに個人に係る情報を含むケースにおいては、このような情報までがどうして必要なのかといった不安を与えないように、丁寧な説明しておくことが求められます。その説明が不十分な場合、データが提供して貰えない、または、不当なデータ収集をしているなどの悪評が広がってしまうことも起こりかねません。

1.2.1 個人に係る情報の有無
目的のために収集するデータには、個人に係る情報を含みますか

4.1.2 人の考えや行動の変化
この派生データや活用方法が、人の考えや行動に変化を与える可能性がありますか



- データ活用において、まず前提たる土台となる「**情報保護**」が必要で、
- その上で「**安全・安心**」を確保した上で、
- 「**公平性**」と「**説明責任**」のあるデータ活用を実現することで、
- このデータ活用の「**正当性**」が示すことができる

倫理	つまずきポイント
1 正当性	
	1.1 活用にあたっての納得感 1.2 個人に係る情報の提供に対する不安対処 1.3 個人に係る情報の提供リスクとメリットのバランス
2 説明責任	
	2.1 派生データの結果に対する説明 2.2 派生データに偏りや間違いの混入の可能性
3 公平性	
	3.1 派生データによる不公平感や差別感の可能性
4 安全・安心	
	4.1 派生データが個人や組織や地域に与える影響の可能性 4.2 悪用の可能性 4.3 国内でのデータの保管/管理
5 情報保護	
	5.1 情報セキュリティ対策 5.2 提供/公開先のコントロール 5.3 収集データの適切な削除 5.4 他データの突合での個人特定 5.5 提供データに関する権利侵害



つまずきポイントと、その有無をチェックするためのリスト 対応策となり得る事項①

倫理	つまずきポイント	チェック概要	
1 正当性	1.1 活用にあたっての納得感	1.1.1 目的の正当性 1.1.2 収集の正当性 1.1.3 生成方法の正当性 1.1.4 種類や意味合いの正当性 1.1.5 活用の仕方の正当性 1.1.6 正当性の説明	
	1.2 個人に係る情報の提供に対する不安対処	1.2.1 個人に係る情報の有無 1.2.2 個人に係る情報の提供の不安感	
	1.3 個人に係る情報の提供リスクとメリットのバランス	1.3.1 プライバシーリスクとメリットのバランス 1.3.2 プライバシー情報提供の選択権	
	2 説明責任	2.1 派生データの結果に対する説明	
		2.2 派生データに偏りや間違いの混入の可能性	
	3 公平性	3.1 派生データによる不公平感や差別感の可能性	
	4 安全・安心	4.1 派生データが個人や組織や地域に与える影響の可能性	
		4.2 悪用の可能性	
		4.3 国内でのデータの保管/管理	
	5 情報保護	5.1 情報セキュリティ対策	
5.2 提供/公開先のコントロール			
5.3 収集データの適切な削除			
5.4 他データの突合での個人特定			
5.5 提供データに関する権利侵害			

	問題事象	リスク要因	リスク低減策
1.1	サービスの利用者がつかない	サービスの意義、個人情報の必要性に納得えられない	判り易く説明 (多くの人のｺﾝｾﾝｽ)
1.2		個人情報が漏洩した時の影響に不安を持たれる	個人情報に対する安全対策の開示
1.3		サービスが開始できなくなる	プライバシーリスクの方が大きいと判断する人もいる
		少数派意見でも、問題として拡散、クローズアップされる	個人情報提供の提供
			リスク対策と共にサービスの意義を丁寧に説明



つまずきポイントと、その有無をチェックするためのリスト 対応策となり得る事項②

倫理	つまずきポイント	チェック概要
1 正当性		
	1.1 活用にあたっての納得感	
	1.2 個人に係る情報の提供に対する不安対処	
	1.3 個人に係る情報の提供リスクとメリットのバランス	
2 説明責任		
	2.1 派生データの結果に対する説明	
	2.1.1 結果に対する納得感	
	2.1.2 結果に対する不信感	
	2.2 派生データに偏りや間違いの混入の可能性	
	2.2.1 結果の偏りや間違い	
	2.2.2 生成で与えるデータの偏り	
	2.2.3 不正データの混入	
	2.2.4 生成方法における限界	
3 公平性		
	3.1 派生データによる不公平感や差別感の可能性	
	3.1.1 善し悪し表現	
	3.1.2 不公平や差別たる指標	
	3.1.3 特定の人にとって不利な提供	
4 安全・安心		
	4.1 派生データが個人や組織や地域に与える影響の可能性	
	4.2 悪用の可能性	
	4.3 国内でのデータの保管/管理	
5 情報保護		
	5.1 情報セキュリティ対策	
	5.2 提供/公開先のコントロール	
	5.3 収集データの適切な削除	
	5.4 他データの突合での個人特定	
	5.5 提供データに関する権利侵害	

	問題事象	リスク要因	リスク低減策
2.1	社会的信用を失う 利用者や社会に問題を与える	AIの分析結果が人には理解できない、その適用で問題発生	推論過程見える化や答え合わせ検証、人による判断
2.2		学習データの収集や作成時に偏りの混入、社会に実在する偏り	学習データの範囲選択/作成手続き適正化、偏り検出ツール活用
3.1		分類や区別する説明したつもりが、差別や偏見と受け止める人もいる	色々な立場の人のいることに配慮した説明
		サービスが利用できない人がいることで不公平との指摘	スコアリングでは算出要素やロジックに透明性を提示
			必要に応じて複数の利用手段の用意



つまずきポイントと、その有無をチェックするためのリスト 対応策となり得る事項③

倫理	つまずきポイント	チェック概要
1 正当性		
	1.1 活用にあたっての納得感	
	1.2 個人に係る情報の提供に対する不安対処	
	1.3 個人に係る情報の提供リスクとメリットのバランス	
2 説明責任		
	2.1 派生データの結果に対する説明	
	2.2 派生データに偏りや間違いの混入の可能性	
3 公平性		
	3.1 派生データによる不公平感や差別感の可能性	
4 安全・安心		
	4.1 派生データが個人や組織や地域に与える影響の可能性	
		4.1.1 個人や組織や地域に悪影響
		4.1.2 人の考えや行動の変化
		4.1.3 人の考えや行動の変化による問題
	4.2 悪用の可能性	
		4.2.1 提供による悪用
		4.2.2 不特定多数への提供
		4.2.3 海外への伝達
	4.3 国内でのデータの保管/管理	
		4.3.1 国外への保管
		4.3.2 クラウドの国外リージョン利用の有無
		4.3.3 データ管理の海外法人への委託
		4.3.4 国外保管禁止情報の有無
5 情報保護		
	5.1 情報セキュリティ対策	
	5.2 提供/公開先のコントロール	
	5.3 収集データの適切な削除	
	5.4 他データの突合での個人特定	
	5.5 提供データに関する権利侵害	

	問題事象	リスク要因	リスク低減策
4.1	社会的信用を失う	発信した情報で風評被害に発展 AIの結果を用いた判断や制御で問題 過去に例のない新たな社会的問題	個人・組織・地域の言及の仕方に配慮 社会インフラ適用に慎重、人間の判断プロセス 人の行動や活動に与える影響の検証
4.2	個人や組織へ迷惑を与える	犯罪の新しい手口に悪用 個人情報悪用は悪用に使われやすい	犯罪に利用される観点での確認 個人情報はより高い漏洩対策
4.3	海外クラウド起因での、個人情報の海外流出	海外リージョンの利用は漏洩リスクが高い 海外クラウドでは外国の当局の要請で国内のデータも提供	個人情報は国内リージョン必須 仮名加工情報活用、日本/海外のクラウドの使い分け

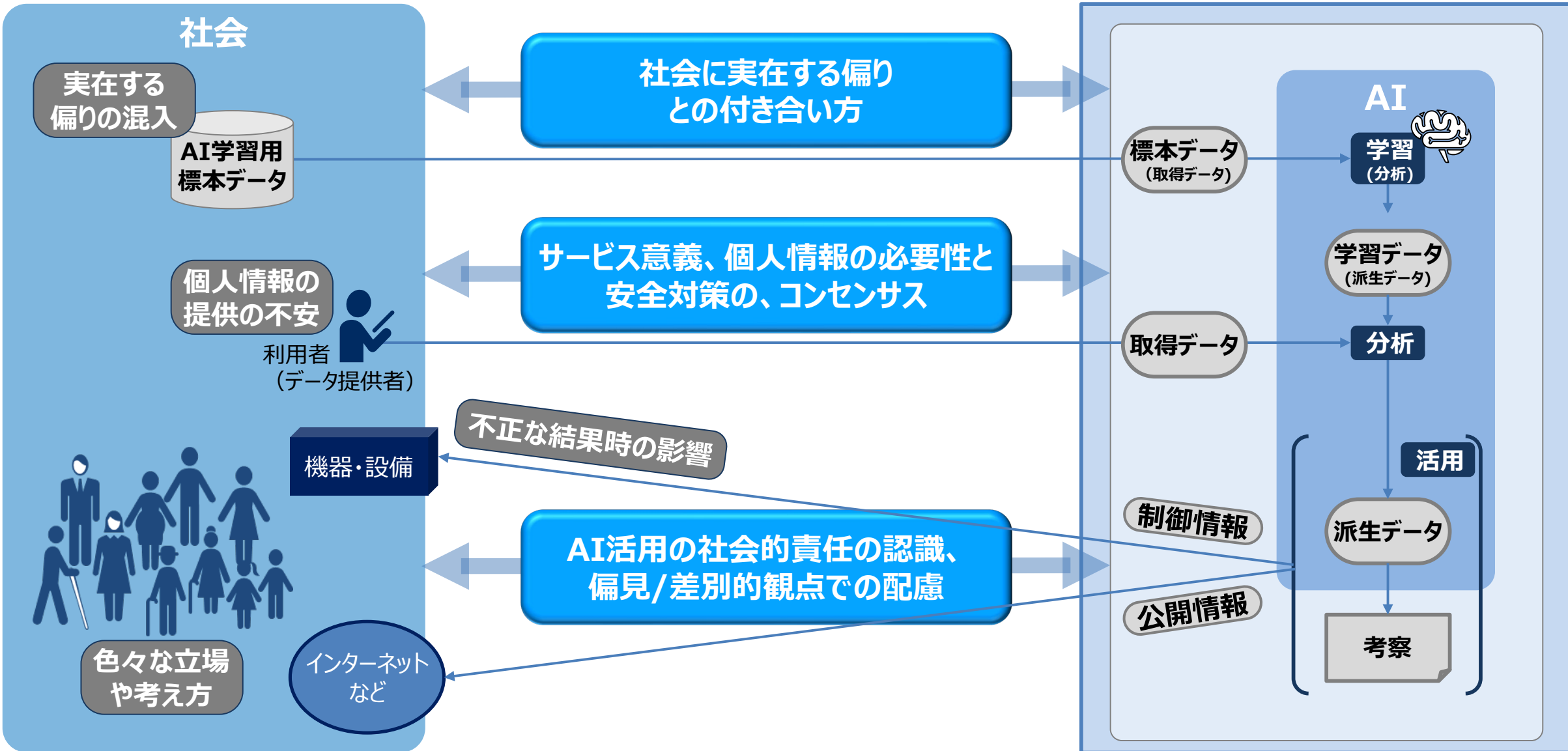


つまずきポイントと、その有無をチェックするためのリスト 対応策となり得る事項④

倫理	つまずきポイント	チェック概要
1 正当性		
2 説明責任		
3 公平性		
4 安全・安心		
5 情報保護		
	5.1 情報セキュリティ対策	5.1.1 機密性対策 5.1.2 完全性対策 5.1.3 可用性対策 5.1.4 対応/取得済の認証規格の有無 5.1.5 対応/取得済の認証規格への対応
	5.2 提供/公開先のコントロール	5.2.1 外部提供 5.2.2 インターネットを通じた提供 5.2.3 提供先の限定 5.2.4 提供先との情報保護の取り決め 5.2.5 国外流出禁止の情報の有無 5.2.6 国外流出禁止に対する対策
	5.3 収集データの適切な削除	5.3.1 個人や組織や地域に係る情報の有無 5.3.2 削除のルールと運用 5.3.3 削除ポリシーの開示
	5.4 他データの突合での個人特定	5.4.1 データの粒度 5.4.2 個人が推測できる特徴の有無 5.4.3 同様な情報の公開サイトの有無 5.4.4 個人コメントの有無
	5.5 提供データに関する権利侵害	5.5.1 価値の高い情報の有無 5.5.2 取得データと異なる価値の有無

	問題事象	リスク要因	リスク低減策
5.1	個人情報の漏洩により、 2次的な被害や社会的信用失墜	一般的なセキュリティリスク	セキュリティ対策、外部認証の取得
5.2		データを扱う組織やシステムの範囲拡大	個人情報扱い制限、仮名加工情報活用
5.3		データ提供先の問題で、個人情報が漏洩	外部組織への提供は仮名/匿名加工情報
5.4	個人の特定による本人に対する問題	データ活用のため、継続保有・利用	個人情報は削除運用、データ活用には仮名加工情報を活用
5.5	データ所有者からの権利侵害の訴え	ネット上の他の情報との突合せが可能	個人の希少性の高い情報の扱い方に配慮
		派生データに対する権利侵害の指摘	元情報に対する同一性や、新しい価値の確認

〔補足1〕 AI・個人情報の活用における「社会との協調たる会話」のイメージ



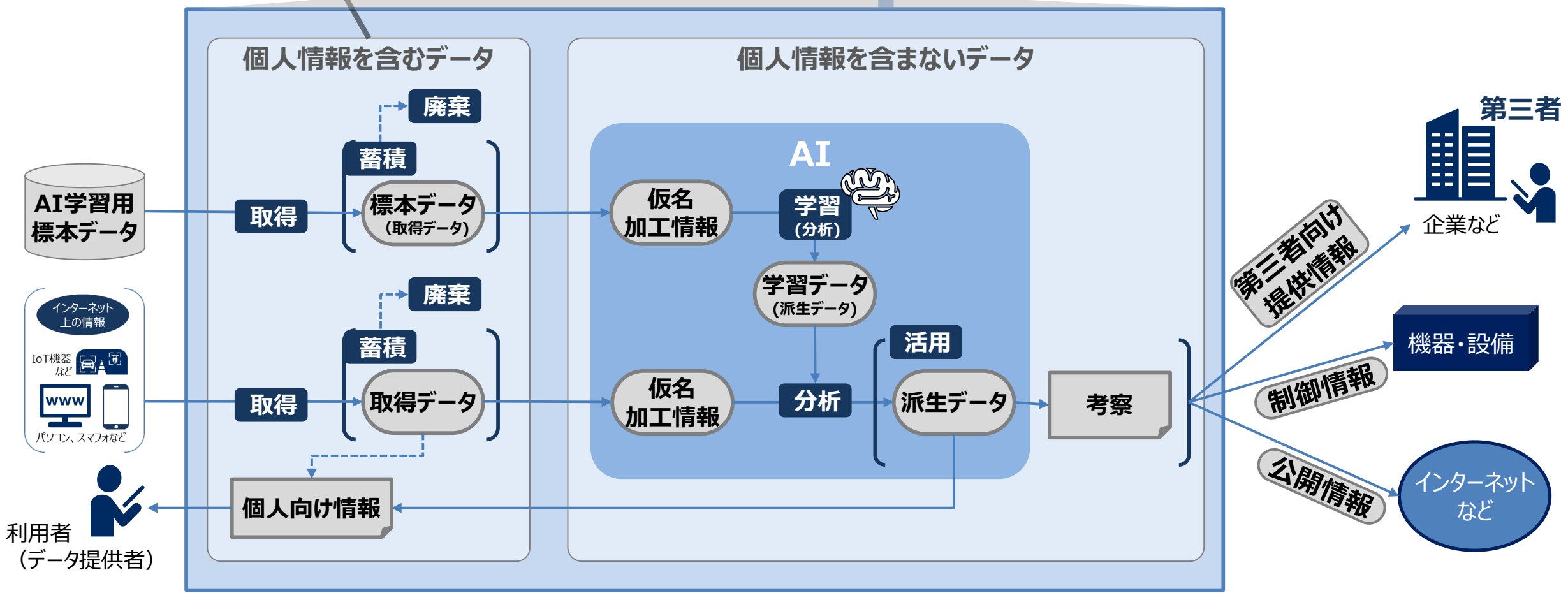
〔補足2〕 AI・個人情報の活用における「漏洩リスク低減への取組み」のイメージ

安全な環境・体制で保管

日本と海外のクラウドの使い分け

共同利用業務委託先

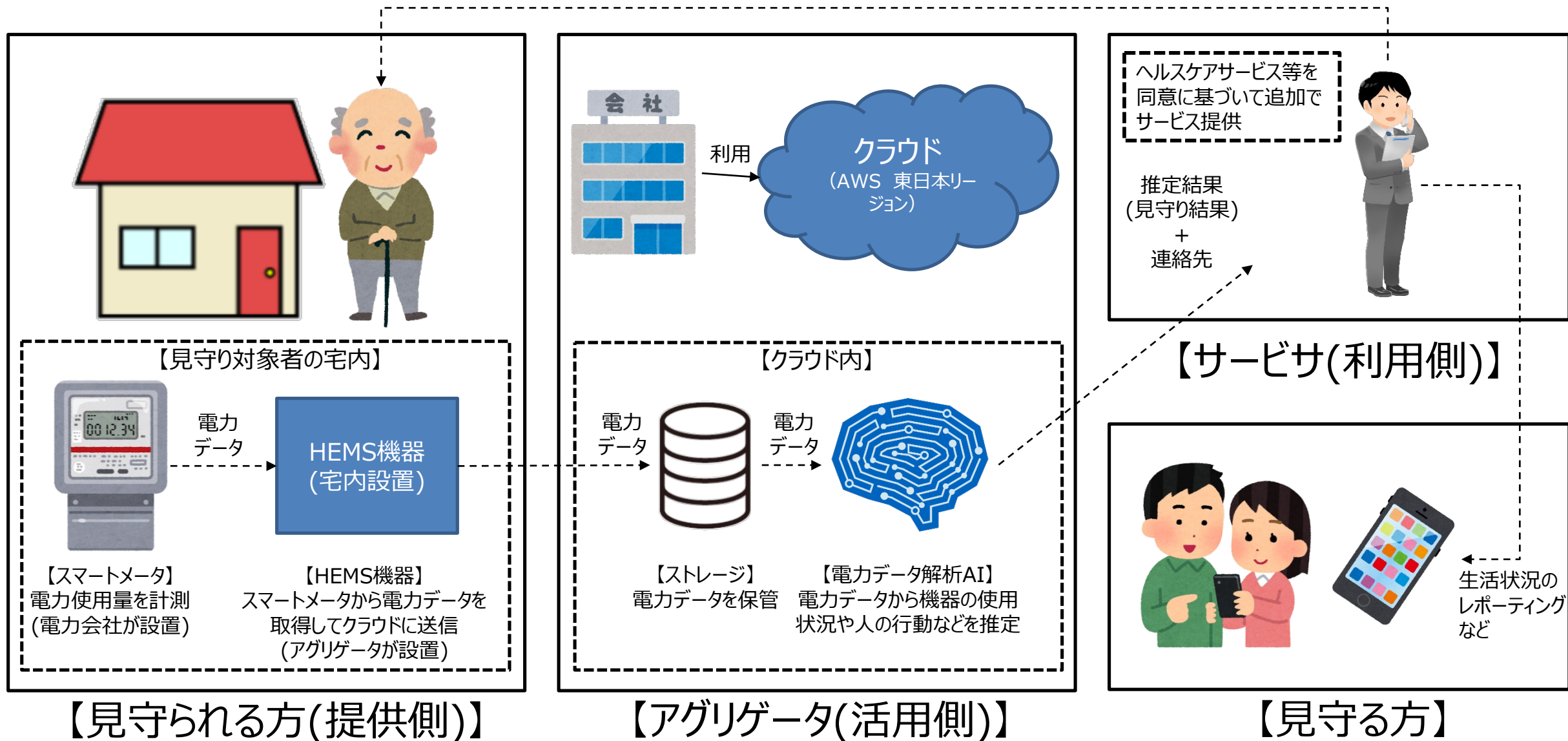
渡すデータのコントロール



研究会の中で検証した事例：スマートメータを利用した“みまもりサービス”



安否確認・ヘルスケアサービスなど



つまずきポイントの洗い出し例 (一部抜粋)



倫理	つまずきポイント	説明	つまずきポイントが潜む部分のチェックリスト	チェック記入欄
1 正当性				
	1.3	個人に係る情報の提供リスクとメリットのバランス		
	1.3.1	個人に係る情報の提供リスクより、メリットの方が大きいといえますか		世帯によってはメリットが大きいと感じられるケースもある。一方で、家族関係等で問題が発生することを考慮して、サービス契約前に、メリット・デメリットの説明が必要。
	1.3.2	リスクとメリットのバランスは、個人の考えに大きく依存しますか(提供選択権の必要性)		個人の考えや個人の状態に大きく依存する。家族関係等に影響することを配慮して、一方的に契約を締結できないように、見守り側と見守られる側双方に配慮が必要。提供については、電力サービスとは異なる分離されたサービスであることを明確にする必要がある。
2 説明責任				
	2.2	派生データに偏りや間違いの混入の可能性		
	2.2.1	得られた派生データの結果に、偏りや間違いの混入の可能性はないですか		電力データという性質から派生データを生成する元データには偏りが無いため、派生データの結果も正確なデータが出る。一方で、様々な観点で偏りが発生する可能性が大きく、派生データの項目について、十分な配慮が必要。また間違いの混入はあり得るため、派生データの影響度合いに配慮し慎重な派生データの生成が必要。
	2.2.2	統計処理や機械学習で与えるデータに偏りはないですか		間違いが混入しなければ原則として現実を反映したデータとなる。一方で、国内データであっても季節や気候、建物の構造などによって偏りが発生するため、データの選び方に工夫が必要。
	2.2.4	派生データの説明に限界がある場合、それを派生データの活用者に説明できていますか		アルゴリズム等の開示が難しい旨は説明しておく。偏りや間違いの混入等についての問題点について説明を丁寧に行う。
4 安全・安心				
	4.1	派生データが個人や組織や地域に与える影響の可能性		
	4.1.1	派生データを提供する場合、個人や組織や地域に悪影響を与える可能性はないですか		不在情報などにより、治安悪化や窃盗というリスクを上昇させる可能性を持っているので、漏洩をしない体制を行うか、個人や世帯と情報が直接的に紐付かないように注意する。派生データ・AIの精度の乱れによって、異常に救急サービスが要請されるされる等による地域社会への負荷増大する可能性がある。
	4.1.2	この派生データや活用方法が、人の考えや行動に変化を与える可能性がありますか		これまでになかったサービス、例えば広告等がサービスから提供されることで、顧客の行動変容がある可能性がある。異常なデータを出さないように、個人が自らの行動を抑制する事による不便や健康被害につながる可能性がある。
	4.1.3	その人の考えや行動の変化が、前例のない問題を引き起こす可能性はないですか		電力データの利用による派生サービスを気にしすぎて、電力利用について配慮しすぎることにより、利用抑制等による不便益を負わないようにする。派生データの生成時のしきい値などにおいて、生活に影響を及ぼさないように、十分な配慮が必要。

- つまづきポイントのチェックリストは、問題になりそうな箇所を明らかにして、サービス・業務を円滑に進めるためのものであり、チェック後には関係者にできるだけ共有を行い、課題解決につなげる必要がある。

■ 活用プロセスの提案

①現状の全体を俯瞰する

- チェックリストの全体を見て、つまづきポイントがどのような問題事象に多いのか把握する。
- 実施しようとしているサービス・業務において大きく影響点が無いか確認する。
- 自社にAIポリシー等がある場合は、再度照らし合わせて大きな課題を明らかにする。

②リスク要因を把握する

- 課題となったつまづきポイントについて、どのようなリスク要因があるかを整理する。
- どの関係者が、どのつまづきポイントに関わっているかを明確にする。

③リスク解消・低減案の策定

- リスクが解消できるか、または低減できるかを整理し対策を考える。法律で明確になっていない範囲が多く、完全に解消することは難しいものが多いので、できるだけ低減する策を、関係者が連携して考える。

➤ つまづきポイントに対して、どのような対応策があるのかを着眼点とともに2021年度の研究会で検討中。



日本データマネジメント・コンソーシアム
Japan Data Management Consortium [JDMC]