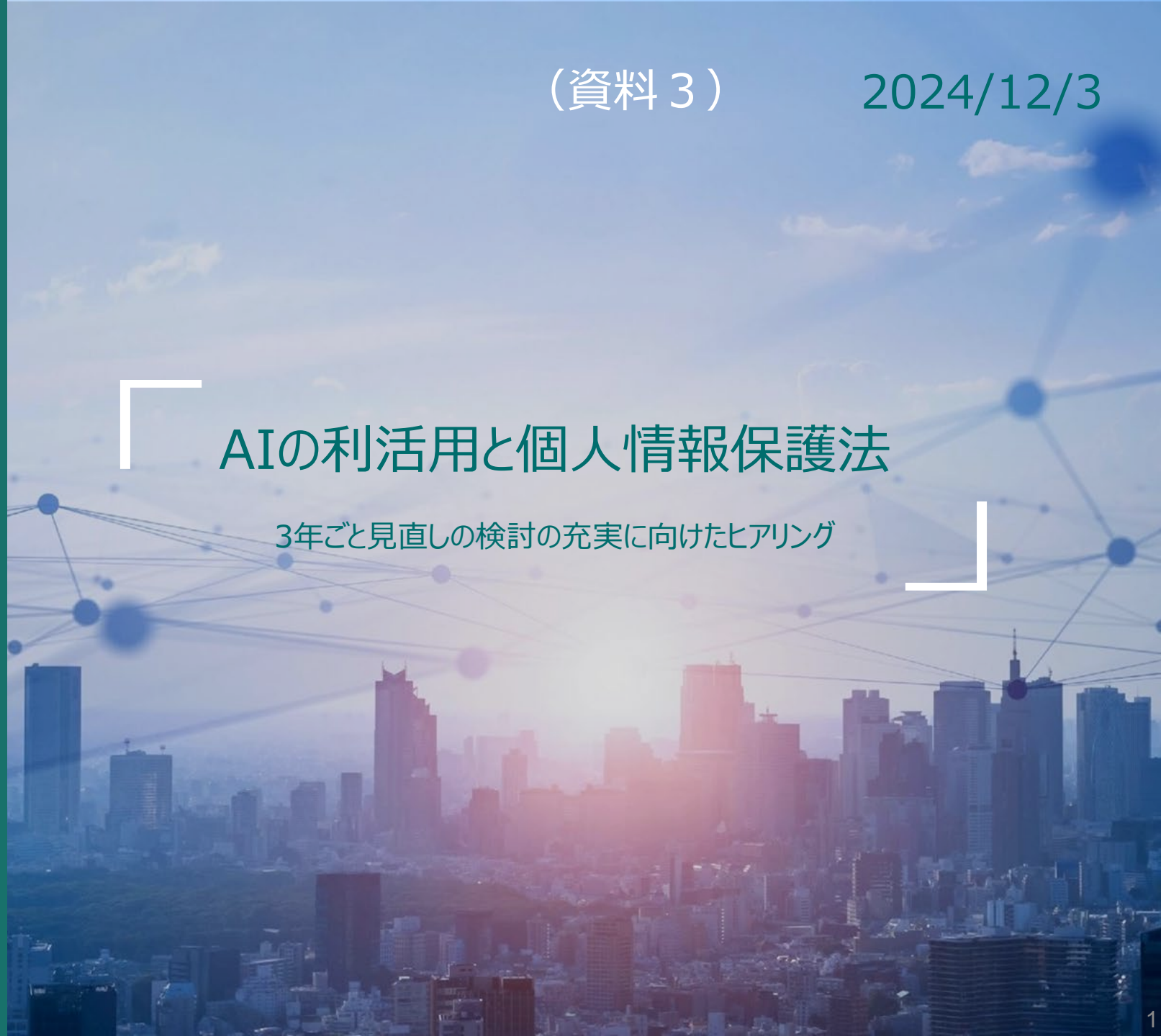
（資料３）　2024/12/3

Japan
Deep Learning
Association

AIの利活用と個人情報保護法

3年ごと見直しの検討の充実に向けたヒアリング



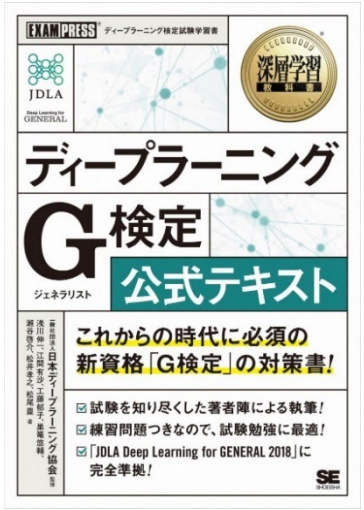
Japan Deep Learning Association

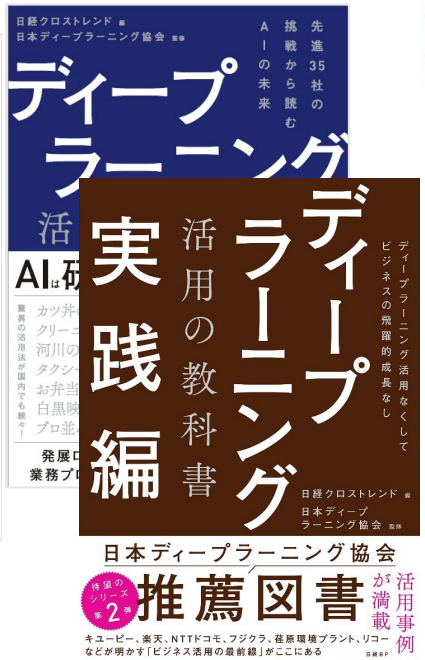
設立目的

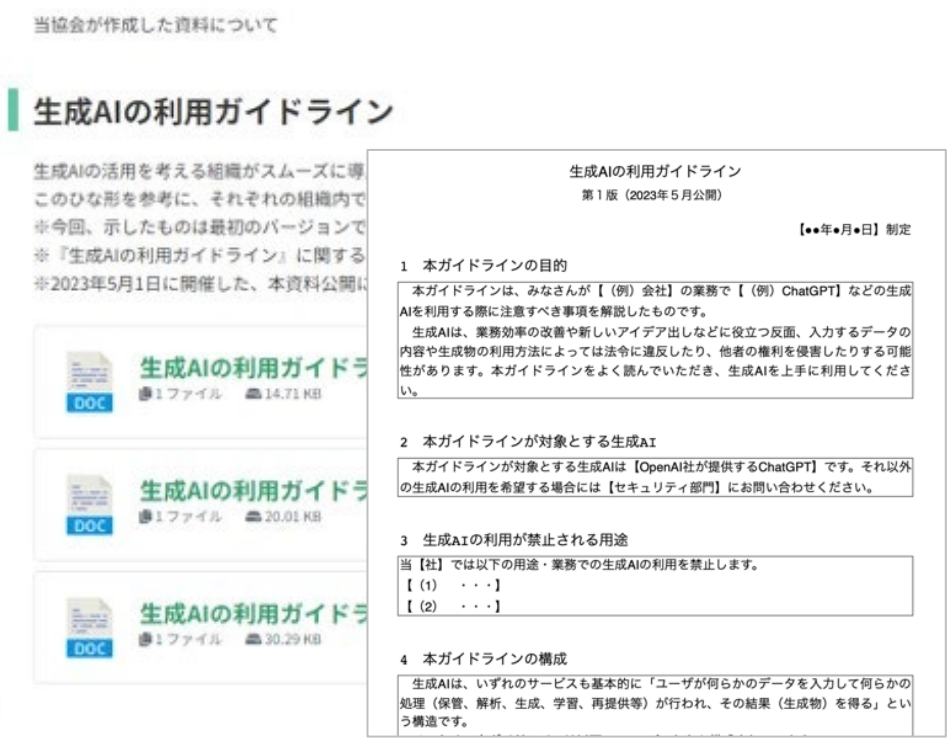
ディープラーニング（DL）技術の活用による日本の産業競争力の向上

・2017年６月設立（理事長：松尾豊）
・DLをビジネスの核とするスタートアップ、研究者を中心とした産学連携団体としてスタート
・**人材育成・活用促進**・社会提言・理解促進・国際連携に取り組む

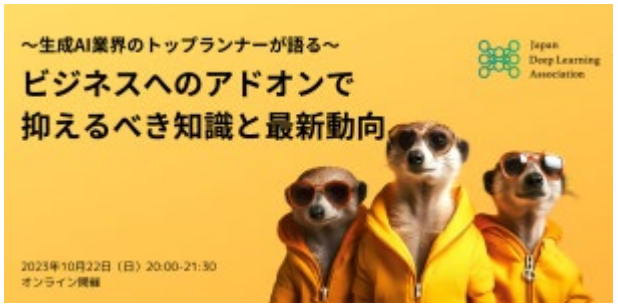
■活動成果物の例







■主催イベントの例



協会概要

名称	一般社団法人　日本ディープラーニング協会
英称	Japan Deep Learning Association　略称：　JDLA

理事長	松尾　豊	東京大学
理事	井﨑　武士	エヌビディア合同会社
	江間　有沙	東京大学国際高等研究所東京カレッジ
	岡崎　直観	東京科学大学　情報理工学院
	岡田　陽介	株式会社ABEJA
	岡田　隆太朗	当協会専務理事
	尾形　哲也	早稲田大学
	柿沼　太一	弁護士法人STORIA法律事務所
	川上　登福	株式会社IGPI Digital Intelligence
	竹川　隆司	株式会社zero to one
	南野　充則	株式会社GROWTH VERSE
	西山　圭太	東京大学未来ビジョン研究センター
	藤吉　弘亘	中部大学
	八木　聡之	富士ソフト株式会社

特別顧問	小宮山　宏	株式会社三菱総合研究所理事長、第28代東京大学総長
	谷口　功	独立行政法人国立高等専門学校機構　理事長
	三村　明夫	日本製鉄株式会社 名誉会長
	森　信親	東京大学大学院経済学研究科金融教育研究センター招聘教授 、元金融庁長官
顧問	森　正弥	株式会社博報堂DYホールディングス CAIO
	吉本　豊	JSR株式会社執行役員
監事	江戸川　泰路	（江戸川公認会計士事務所）

正会員（47社）

ABEJA	Aidemy	AiHUB	AI inside	ALGOMATIC	a|+	AnyTech
Appen	AVILEN	BrainPad	BytePlus	株式会社 調和技研	connectome design	DEEPCORE
deep instinct	EDGE Technology	Elith	FastLabel	FiNC Technologies	GAUSS	GRID
HEROZ	AI Factory	IGPI Digital Intelligence	iXs	キカガク KIKAGAKU	Kaozyt	KUNO
Liaro	MILIZE	morpho	MUSASHi AI	NEURAL	NVIDIA	GPU EATER
OpenFashion	PKSHA TECHNOLOGY	Preferred Networks	Ridge-i	Rist	ROSSO	SECURE
スキルアップNeXt	Spiral.AI	tiwaki Always be better	Weights & Biases	zero one		

上場企業（うち入会後上場13社）

2024年11月現在

有識者会員（21名）

石川 冬樹 国立情報学研究所 アーキテクチャ科学研究系 准教授	**江間 有沙** 東京大学 国際高等研究所東京カレッジ 准教授
牛久 祥孝 オムロン サイニックエックス株式会社 リサーチバイスプレジデント	**岡崎 直観** 東京科学大学 情報理工学院 教授
岡谷 貴之 東北大学 大学院 情報科学研究科 教授	**尾形 哲也** 早稲田大学 基幹理工学部表現工学科 教授
柿沼太一 弁護士法人STORIA法律事務所　代表パートナー弁護士	**北野 宏明** ソニーコンピュータサイエンス研究所 代表取締役社長兼所長
柴山 吉報 阿部・井窪・ 片山法律事務所　弁護士	**巣籠 悠輔** 東京大学大学院 工学系研究科 招聘講師
清田 純 国立研究開発法人理化学研究所 情報統合本部 チームリーダー	**Shane GU** Google DeepMind 研究員
中島 秀之 札幌市立大学 理事長・学長	**羽深 宏樹** 京都大学 法政策共同研究センター 特任教授
藤吉 弘亘 中部大学 工学部 ロボット理工学科 教授	**松尾　豊** 東京大学大学院　工学系研究科　教授
松原 仁 京都橘大学工学部 教授、情報学教育研究センター長	**馬渕 邦美** 一般社団法人Metaverse Japan 共同代表理事
丸山 宏 Preferred Networks PFNシニアアドバイザー	**山下　隆義** 中部大学　工学部　情報工学科　教授
安田孝美 名古屋大学　大学院情報学研究科・情報学部　教授	

2024年11月現在

賛助会員（53社）

PLATINUM

Deloitte. デロイト トーマツ
Deloitte. Digital
baycurrent
KPMG

GOLD

FUJISOFT ＊
MARUI GROUP
SMBC 三井住友銀行
NOMURA
HUAWEI
Adire アディーレ法律事務所
HAIInstitute

SILVER

nico
JTEKT
FUTURE
TAIYO YUDEN
NTT docomo
Stitch
HITACHI 日立システムズ
TOP ENGINEERING INC.
Western Digital.
CIC CREDIT INFORMATION CENTER 株式会社 シー・アイ・シー
KDDI
Orchestrating a brighter world NEC
すべての革新は患者さんのために CHUGAI 中外製薬 ロシュ グループ
THIRDWAVE
AGAROOT
第一工業製薬
開志専門職大学 KAISHI PROFESSIONAL UNIVERSITY
BRYCEN Co.,Ltd.
Motivational Engineer CLINKS
アイリスオーヤマ
ABeam Consulting
APC All Personal Property Center
Access net Access your smile Access to the future
TOPY TOPY INDUSTRIES, LIMITED
株式会社ジェイ・ウィル・エックス
HITACHI Inspire the Next 日立システムズエンジニアリングサービス
Jeki Data-Driven Lab
Google
MOTEX
GMO MEDIA
aws
AKKODiS
MITANI SANGYO CO.,LTD.
NRI SECURE
GMO INTERNET GROUP
pwc
九州電力
株式会社ベネッセコーポレーション
GDEP Solutions
Smiley
SI&C
三菱商事株式会社
株式会社 KOIRASE

＊事務局出向

2024年11月現在

行政会員（28団体）

地方公共団体

札幌市 City of Sapporo 北海道 札幌市	仙台市 SENDAI CITY 宮城県 仙台市	 栃木県 那須塩原市	 東京県 文京区	新潟県 新潟県
 新潟県 長岡市	 長野県 塩尻市	 長野県 松本市	豊田市 Toyota City 愛知県 豊田市	 愛知県 名古屋市
 石川県 加賀市	三重県 伊勢市	 滋賀県	和歌山県 Wakayama Prefecture 和歌山県	 岡山県津山市
 広島県	 島根県	 山口県	 三豊市 香川県 三豊市	高知県 Kochi Prefecture 高知県
 香川県坂出市	徳島県	 大分県大分市	 おんせん県おおいた 大分県	 福岡県 北九州市
 宮崎県 都城市				
 広島県教育委員会	山口県教育委員会			

教育委員会

2024年10月現在

目次

要配慮個人情報の収集

検討の方向性

関連項目

- ２　個別の個人の権利利益への直接的な影響が想定されない個人データの利用に対する規律の考え方

要配慮個人情報の取得についての規律を、データベース等を構成するようなもののみに限定するか、少なくとも生成AIの学習目的のような一般的・汎用的な分析結果の獲得と利用のみ目的とするものを除外するべきである。

想定するケース

Common Crawlが公開しているデータセットを加工したデータや、自らクローリングによりWebサイトから収集したデータを用いてデータセットを作成し、当該データセットを用いて学習した学習済みモデルとデータセットを公開するようなケース。

現状

- あらかじめ本人の同意を得ないで要配慮個人情報を取得することを原則として禁止（法20条2項）。
- 「OpenAI に対する注意喚起の概要」において、①収集する情報に要配慮個人情報が含まれないよう必要な取組を行うこと、② 情報の収集後できる限り即時に、収集した情報に含まれ得る要配慮個人情報をできる限り減少させるための措置を講ずること等が求められている。

要配慮個人情報の収集

個人の権利利益への影響

- そもそも、Webサイト上に存在する要配慮個人情報の大半は、本人や報道機関等により掲載されたものであり、取得に際して本人同意が不要なものである。例えば、Common Crawlが公開しているデータセットの中で、病名または犯罪名が含まれるテキストデータ8,665件を目視により確認したところ、本人の同意なく取得することが許されない可能性のあるデータはわずか3件であった。
- Webサイト上で取得可能な要配慮個人情報のうち、本人の同意なく取得できない可能性のある「要配慮個人情報」を類型化すると、①親族等の病歴について記載されているもの、②誹謗中傷として社会的身分や人種等を用いているもの、③少年犯罪に関して実名が記載されているもの、といったものに限られるように思われる。しかしながら、このような情報だけを選別して削除することは極めて困難であり、「収集した情報に含まれ得る要配慮個人情報をできる限り減少」させたとしても削除できる保証はない。なお、保守的な対応をして過剰にフィルタリングをしてしまった場合、病気や犯罪等についての学習データが不足することになり、モデルの性能を大きく損なう可能性がある。
- クローリングにより取得するデータは、一般に使用されるPCやソフトウェアでは利用できない形式・規模等のものであり、収集・公開により当該データが人の目につくことによる権利利益への影響はほとんどない。
- そもそも、Webサイト上のデータはCommon Crawlによって収集・公開されており、これをさらに公開することによる個人の権利利益への影響はほとんどない。

要配慮個人情報の収集

必要性

- 生成AIの開発においては、学習用データの規模が非常に重要であり、クローリング等によりWebサイト上に公開されているデータを収集する必要性が高い。実際には、Common Crawlにより収集・公開されたデータセットを加工して用いることも広く行われている。
- 生成AIの開発に投入できる金額が限られている日本企業では、個社で海外のビッグテックに対抗することは難しく、技術のオープン化が重要になりうる。基盤モデルを開発しオープンにする場合、基盤モデルだけでなく、データセットも併せて公開することが有益である。

現状の課題

- クローリングの技術自体は検索エンジン等にも広く用いられてきたものである。Common Crawlは2007年からデータの収集・公表を行っている＊。
- Common Crawlのデータを利用する場合、一括でダウンロードするデータについて「収集する情報に要配慮個人情報が含まれないよう必要な取組を行うこと」は困難である。
- 海外のビッグテックが、海外において、自社内で日本語Webサイトのデータを収集する場合に、すべての企業が注意喚起を遵守するかは不明であり、現状の規律の実効性に疑問がある。他方、日本企業がデータを収集・公表する場合には、「できる限り」といった注意喚起の文言もあいまって萎縮効果を及ぼす。

＊ Common CrawlのWebページ（https://commoncrawl.org/ ）参照。

委託を受けたデータを用いた学習等

検討の方向性

関連項目
- ２　個別の個人の権利利益への直接的な影響が想定されない個人データの利用に対する規律の考え方
- ３　個人データの第三者提供を原則として禁止する仕組みの妥当性

AIの学習は、統計処理と同様、一般的・汎用的な分析結果の獲得と利用のみを目的とするものであり、委託や利用目的の通知等において統計処理と同様に扱うとともに、このような利用について第三者提供の規律を見直すべきである。

想定するケース

一例として、顔画像や医療データ等の個人データの入力・解析を行うAIを組み込んだSaaSを提供する企業において、当該SaaSのユーザーである企業（個人データの委託元）が入力したデータをAIの学習に用いるようなケース。学習に用いる場合、①ユーザー企業1社のデータを学習するパターン、②複数のユーザー企業のデータを合わせて（個人データの突合はせずに）学習するパターン、③複数のユーザー企業のデータを、個人データを突合したうえで学習するパターンが考えられる。

委託を受けたデータを用いた学習等

必要性

- AIの学習は一般的・汎用的な分析結果の獲得と利用のみを目的とする点でまさに統計処理と同視しうるものであり、個人情報保護法の規律としても異なる扱いをすべきではない。
- 個人情報を解析するSaaSなどでは、学習や解析対象のデータとして、複数のサービスユーザーから受領した個人情報を1つの学習用データセット（データベース）として学習・情報解析に供する必要性が高く、「委託」等の枠組みにおいて、かかるデータの利用ができることを明確化する必要性が高い。
- 複数社から受領した個人データを突合したうえで利用することについても、PETsの活用等により一定の範囲で本人の同意なく認める必要性は高い。
- 企業が持つ有用なデータが、AIの技術の発展により取得時に想定していなかった分野のAIの学習に用いることもありうる。

個人の権利利益への影響

- AIの学習は基本的に一般的・汎用的な分析結果の獲得と利用のみを目的とするものであり、個人の権利利益への影響は低い（意図的な過学習といった一般的・汎用的な分析結果の獲得以外を目的とする不適正な利用は除く。）
- 個人の権利利益への影響において重要なのは利用目的の通知等ではなく、安全性確保のための取組みや学習以外に用いられないことの担保である。

生成AIの利用の局面における規律

関連項目

- ３ 個人データの第三者提供を原則として禁止する仕組みの妥当性
- ４ 個人データの取扱い態様の多様化の下における、データの適正な取扱いに係る義務を負うべき者の在り方

検討の方向性

生成AIの利用の局面において、いわゆるクラウド例外（またはそれに準じた例外）、委託及び越境移転についての適用範囲や要件の明確化を行うべきである。また、サービス利用者のみに責任を負わせるべきではない。

必要性

- 生成AIを含む多くのSaaS等のサービスは、サービス提供にあたり複数の事業者が関与するものとなっているが、これらのサービスの利用において、現状のクラウド例外や委託の規律だけでは対応が難しい。特に、一つのサービス内で複数の生成AIを利用できるサービスを利用する場合、利用する企業はすべての生成AIについて利用規約等を確認してサーバーの所在地や個人データの取扱いについて調査をする必要があるが、具体的なケースにおけるクラウド例外の該当性や、委託先の監督等の義務、及び越境移転に際しての基準適合体制の確認・判断は難しく、かつ利用者側に過大な負担となっている。
- 少なくとも、サービスの利用者において個人データを入力する場合の法的な枠組みとその要件を明確にしたうえで、利用者において要件該当性の判断のための情報を容易に取得できるようにする必要がある。
- サービス利用者が提供元基準や容易照合性を理解していないことに起因して、認識なく個人データを入力しているケース（サービス提供者が認識なく個人データを受領しているケース）も散見される。

今後の個人情報保護法制の検討

- EHDSなど、欧州等でのデータの利活用が進んでいる現状を踏まえ、ドメインごとの特別法等による利活用の促進を見据えた議論も必要である。
- AI等の新しい技術について、「問題が起きる前」に検討することで、個人の権利利益に配慮した技術の利用促進につながる。
- 新しい技術に対応するためには、透明性と対話を通じた各ステークホルダーの連携、自主規制やソフトローとの連続性の確保等も必要である。