※本文の複製、転載、改変、再配布を禁止します。

論　説

AIビジネスの最前線からお送りする「AI・データの利用に関する契約ガイドライン（AI編）」の概説

柿　沼　太　一*

抄　録　2018年６月15日，経済産業省によって「AI・データの利用に関する契約ガイドライン[1)]」が策定された。この「AI・データの利用に関する契約ガイドライン」の「AI編」（以下AIガイドラインという）においては「データを保有しているユーザがAI開発技術を有しているAIベンダに対して，保有データを利用したAIの開発を委託する」という典型的な開発パターンにおける，契約締結交渉ポイントや具体的な契約条項について紹介をしている。本稿においては，AIガイドラインをもとに，通常のシステム開発とAIソフトウェア開発の違いや，それに伴う契約条項の留意点を解説する。また，解説に加えて，通常のシステム開発では可能であったベンダによる性能保証やベンダの損害賠償責任の定め等について，AI開発契約においては契約実務上どのような対応が可能なのかについても提案・紹介する[2)]。

目　次

1．はじめに

筆者は近時AI開発をベンダに発注しようとするユーザや，ユーザからデータの提供を受けてAI開発を受託しようとするAIベンダ双方の立場からご相談を頂くが，相談内容は多岐にわたる。

たとえば以下のようなご相談である。

・ユーザから学習済みモデルを用いた出力の精度について一定の保証をするようにと強く要請された場合，AIベンダはどのように対応すべきか

・AIベンダに学習済みモデルの開発を発注するに際して，AIベンダから性能保証ができ

* 弁護士　Taichi KAKINUMA

※本文の複製、転載、改変、再配布を禁止します。

ないと言われたが，性能保証ができないのに高額な開発費用を支払うという点に不安があるがどのようにリスクヘッジしたらよいか
・開発成果（学習用データセットや学習済みモデルなど）に関する権利や知的財産権に関してどのようなポイントに着目して交渉し，どのような契約条項に落とし込んだらよいのか
・ユーザの提供データを用いてベンダが開発したAIを組み込んだシステムが誤作動をしてユーザや第三者に損害を与えた場合，誰が責任を負うのか
・AIの誤作動に備えて，AI開発契約においてはどのような定めをすべきか

また，相談を受けるAI利用分野も様々であり，医療画像上の病変を自動的に検出する医療画像AI，疾病推測AI，自動運転車用AI，工場内で半製品・完成品を判定するAIなどがある。

これらAI開発契約に関する相談を分類すると，ほぼ以下の３つの領域のどこかに当てはまる。
・性能保証，検収，瑕疵担保

AI開発契約において「性能保証」「検収」「瑕疵担保」についてはどのように定めればいいのか
・権利・知財

生成された学習用データセット，学習済みモデル，学習済みパラメータは誰がどのような権利を持っているのか
・責任

AI開発・利用に際して生じる可能性のある損害についてAI開発契約ではどのように定めたらよいか

そして，これら現場の悩みは一言で言うとAIソフトウェア開発と通常のルールベースのソフトウェア開発との相違に由来するため，本稿ではまずその点について解説した後に，「性能保証，検収，瑕疵担保」「権利・知財」「責任」の順序で解説する。

なお，「AI」という用語の意味は多義的であるが，AIガイドラインでは「AI技術」のことを「『機械学習』，またはそれに関連する一連のソフトウェア技術のいずれか」という意味で使っているため，本記事では「AI」という用語を「（統計的）機械学習」という意味で使っている。

また，本記事中の各用語の定義は，特別の断りがない限り，AIガイドラインでの各用語の定義に従っている。

2．性能保証，検収，瑕疵担保，について

2．1　通常のシステム開発とAIソフトウェア開発の違い

AI技術を用いない通常のシステム開発契約においては，成果物の性能についてベンダが一定の性能を保証したり，成果物について一定の検収基準に従った検収をユーザが行って合否の判定を行ったり，成果物について瑕疵が存在した場合の瑕疵担保責任を定めるのが通常である。

そのため，AIソフトウェアの開発契約においても，ユーザはベンダに対して一定の性能保証，検収規定や瑕疵担保条項を盛り込むように要求するケースが多く見られる。

しかしAIソフトウェアの場合は，その技術的特性から以下の特徴を有しているため，ベンダがユーザの性能保証などの要求に応じられず，ベンダとユーザの歩み寄りが不可能となってしまうケースがある。
・訓練データに統計的バイアスが含まれることが避けられないため，未知のデータに対する性能保証は原理的に困難
・何か不具合が生じた場合でも，その原因（データの品質，ハイパーパラメータ設定，ソースコードのバグ等）が複数存在し，問題の切り分けが困難
・成果物の検収に際して学習に利用しない独立

※本文の複製、転載、改変、再配布を禁止します。

したデータセットが必要であり，かつ当然のことだが未知データでのテストが不可能

AIソフトウェア開発において「性能保証」，「検収」，「瑕疵担保」に関する双方の意見が対立する原因は，通常のシステム開発が「演繹的」な開発手法であるのに対して，AIソフトウェア開発が「帰納的」である点があげられる。

「演繹」とは，一般的な前提やルールから結論を得る考え方であり「A＝B」「B＝C」「よってA＝C」という三段論法はその典型例とされている。

一方「帰納」とは，複数の個別事例や経験則などの前提を集めて，そこから普遍的な法則を見いだす考え方である。

通常のシステム開発は，まず仕様を確定したうえで，その仕様を実現するためにはこうする必要がある，ということを積み重ねてシステムを開発するという演繹的な開発手法をとる。

一方，AIソフトウェア開発は，開発目的との関係で意味のあると思われる大量のデータを集めてきて，それを用いて学習をさせ，当該データに共通する法則・特徴を見つけ出すという帰納的な開発手法をとる。

帰納的な開発手法により開発されたAIソフトウェアは，理屈を積み上げて開発したわけではないため，学習に使っていないデータを入力した場合，どのような挙動をするかが予測困難であり，「未知のデータでの性能保証が困難」，「なにをもって『瑕疵』というかはっきりしない」ということになる。

2. 2 「通常のシステム開発とAIソフトウェア開発の違い」を乗り越える３つのポイント

このような，「通常のシステム開発とAIソフトウェア開発の違い」をAI開発契約締結に際して乗り越えるためには，筆者は（1）AI開発の特性をユーザとベンダが理解すること，（2）開発プロセスおよび契約の分割，（3）契約内容の工夫，の３点が重要であると考えている。

(1) AI開発の特性をユーザとベンダが理解すること

通常のシステム開発とAIソフトウェア開発においては，前述のように開発手法の発想が異なるが，その点についてユーザ・ベンダが共通認識を持つことが非常に重要である。

AIガイドラインも，この点を強く意識して作られた。このガイドラインが，かなりの分量を割いてAI技術について記述している（「第２AI技術の解説」）のはそのためである。

ユーザとベンダの契約交渉が煮詰まる前に，法務・知財部門の方々はAIガイドラインを参考として，AIソフトウェア開発の特性を理解いただければ幸いである。

(2) 開発プロセスおよび契約の分割

次に重要なのが「開発プロセスおよび契約の分割」である。

AIソフトウェア開発の特徴を非常に乱暴に言ってしまうと「どのようなものができあがるか事前に予測することがユーザ・ベンダ双方にとって困難であること」，要するに「開発を進めてみないと，うまくいくかどうかわからない」という点にある。

この「うまくいくかどうかわからない」というのは「ベンダはわかっているがユーザはわからない」という情報の非対称性の話ではなく，「AIソフトウェアの原理上ベンダもユーザもわからない」ということを意味している。このような特徴はユーザ・ベンダ双方にとって大きなリスクとなる。

このリスクをコントロールするための一つの方法として，AIガイドラインで提唱しているのが開発プロセスおよび契約を分割する「探索的段階型」の開発手法である。

※本文の複製、転載、改変、再配布を禁止します。

具体的には，AIソフトウェアの開発を，①アセスメント段階，②PoC段階[3)]，③開発段階，④追加学習段階の４段階に分けた開発手法である。

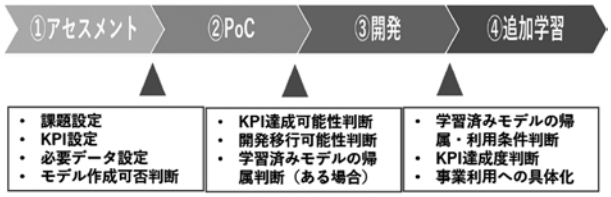

図１　開発プロセスの分割事例

各段階の目的・成果物・契約内容の概要は以下の表１のとおりである。

表１　各段階の目的・成果物・契約内容の概要

	アセスメント	PoC	開発	追加学習
目的	一定量のデータを用いて学習済みモデルの生成可能性を検証する	学習用データセットを用いてユーザが希望する精度の学習済みモデルが生成できるかどうかを検証する	学習済みモデルを生成する	ベンダが納品した学習済みモデルについて、追加の学習用データセットを使って学習をする
成果物	レポート等	レポート等	学習済みモデル等	再利用モデル等
契約	秘密保持契約書等	導入検証契約書等	ソフトウェア開発契約書	例：保守契約、学習支援契約または別途新たなソフトウェア開発契約など

もちろん，「４つ」という数字に意味がある訳ではなく，「うまくいくかどうかわからないから少しつずつ進め，うまくいきそうなら次のステップに行き，無理そうなら中止する。それによってユーザ・ベンダ双方のリスクをコントロールする」という点が本質である。

そのため，開発規模によっては，アセスメント段階とPoC段階が一体となることもあるし，PoC段階を更に複数に分割することもある。

なお，このようにPoC段階と開発段階で契約を分けるとベンダにとっては，「PoC段階で結果が出たにもかかわらず開発段階に移行しない」，あるいは「PoC段階で納品した学習済みモデルについて開発段階に移行しないにもかかわらずユーザによって異なるベンダに横展開される」というリスクが生じる。

前者については，一般的にPoC段階はベンダにとって大きな手間がかかり，開発段階に移行することで初めて投下コストを回収できるという面があるため，ベンダとしては切実な問題である。PoC段階と開発段階とで契約を分割する以上，法的な拘束力を持つ移行義務をユーザに課すことは困難だが，契約上移行に向けての協議義務を課すことが考えられる。

後者については言うまでもなくベンダにとって非常に大きなリスクである。PoC段階で学習済みモデルを提供するのであれば，あくまでモデルの検証目的の提供であること，PoC段階のみで契約が終了した場合にはモデルの廃棄・返還義務を契約上明記しておく必要がある。

(3) 契約内容の工夫

通常のシステム開発とAIソフトウェア開発における契約内容を対比すると以下の表２のようになる。

表２　通常のシステム開発とAIソフトウェア開発契約の対比

	通常のシステム開発	AI開発
契約の法的性質	工程によって異なる（上流工程は準委任型、下流に行くにしたがって請負型）	全工程で準委任型が親和的
完成義務	請負型が適用される工程では完成義務あり	なし（モデル開発契約７条）。ただし成果完成型の準委任契約も締結可能。
性能保証	請負型が適用される工程では合意可能	なし（モデル開発契約7条）。ただし、一定の既知データを用いた場合の性能であれば保証可能な場合もある。
瑕疵担保責任	請負型が適用される工程においては瑕疵担保責任あり	なし

筆者はセミナーなどでこの表２を使って説明することが多いが，そこで一番聞かれる質問は「現場にAIソフトウェアを投入し，未知のデータが入力された場合に性能保証ができないのは

※本文の複製、転載、改変、再配布を禁止します。

理屈としてはわかった。しかし委託料を支払って開発したAIソフトウェアについて一切性能保証がない，というのはやはりユーザとしては納得し難い。契約上の工夫でなんとかできないか」というものである。

この点について実際に提案している方法は「学習に利用しない一定の既知の評価用データを利用した性能保証を行う方法」と「準委任契約における成果完成型を選択する方法」である。

前者については，未知データ入力の場合の保証は難しいとしても，既知の評価用データ入力の場合の保証は技術的には可能なため，選択肢の１つとすることが可能である。

もっとも，このような合意をする場合，当該評価用データが，AIを現場に投入した場合の実際の未知データの性質を十分に反映している必要がある。そのため，その点についてユーザとベンダどちらが保証をするのかが困難な問題として残る。

後者について，まず，「準委任契約」とは，仕事の完成ではなく，一定の事務処理行為を行うことを約する契約のことを言い，請負人が仕事を完成することを約する「請負契約」と対比されることが多い。

AI開発の場合は，そもそもその技術の特性上「仕様確定」「精度保証」「完成」が観念しがたいという特色があるため，仕事完成を約する「請負契約」ではなく「準委任契約」が親和的であると思われる。

さらに，準委任契約には，委任事務の履行により得られる成果に対して報酬を支払うことを約する「成果完成型」と，委任事務の処理の割合に応じて報酬を支払う「履行割合型」があると言われている。

「履行割合型」の典型例は，システム開発契約におけるいわゆる「人月単価方式」である。一定の技術レベルの人員が一定期間稼働することに対して一定の対価を支払う方式であり，成果物の精度や質が委託料の金額や支払い方法に連動しない。

一方，「成果完成型」とはたとえば，弁護士が訴訟事件について依頼者と委任契約を締結する際に，一定の成果に応じて報酬を受領する旨を約する（いわゆる成功報酬制）方式などが該当する。

この「成果完成型」の準委任契約であれば，進捗に応じた成果（たとえば既知データに対する一定レベルの性能など）に対して固定金額を支払う，あるいは成果に満たない場合には一定額の返金をするなどの合意をすることになり[4)]成果未達成の場合のユーザ・ベンダ双方のリスクを下げることができると考えられる。

2. 3 小　括

以上，AIソフトウェア開発において交渉が難航することが多い「AIソフトウェアの性能保証，検収，瑕疵担保」が「通常のシステム開発とAIソフトウェア開発の違い」に起因すること，及びそれを乗り越えるための実務的な工夫について述べた。

なお，AIソフトウェア開発案件において，当初ユーザサイドのビジネス部門や技術部門とベンダとが盛り上がり，その話がユーザサイドの法務・知財部門に上がった瞬間にストップがかかるケースが多く見られる。

そのようなときには，本記事で解説したようなポイントに双方が留意し，合理的な契約交渉が進むことを期待したい。

3． 権利・知財について

AIソフトウェアの開発に際して，成果物等に関する権利・知財について契約上どのように定めるかはユーザ・ベンダにとって非常に大きな関心事である。

典型的な相談は「AIベンダとして，事業会社との共同プロジェクトを立ち上げる際の，知

※本文の複製、転載、改変、再配布を禁止します。

財・法務に対する最初のニギリ方（契約書への文言の入れ方）が事業会社側，AIベンダ側とも曖昧なケースが多く，実際に売り上げが発生した際にモメそうな不安がある。」というようなものである。

3. 1　なぜ交渉が難航するのか

AIソフトウェアの開発においては，通常のシステム開発以上に成果物の権利・知財に関する当事者双方の主張の対立が先鋭化することが多いが，その理由は以下の２点にあると思われる。

・通常のシステム開発と異なり，AIソフトウェア開発においては複数の材料，中間成果物，成果物が存在する
・開発に要する材料，中間成果物，成果物が高い価値を持ち，ユーザ・ベンダ共にそれらを独占／再利用したいという需要が存在する

まず，通常のシステム開発の開発工程をごくごく単純化すると図２に表すように，「『ベンダの労力やノウハウ』という材料を投入して『プログラムや文書類』という成果物を開発する」ということになる。

図２　通常のシステム開発の開発工程

一方，AIソフトウェア開発の開発工程は，図３に表すように「『生データ，学習用プログラム，労力，ノウハウ』という材料を投入することで『学習済みモデル』や「学習済みパラメータ』という成果物を開発し，開発の過程で『学習用データセット』や『発明，ノウハウ』といった中間的な成果物が生じる」という点が大きな特徴である[5)]。

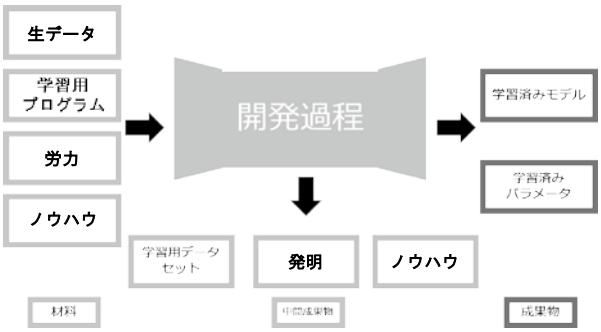

図３　AIソフトウェア開発の開発工程

さらに，これらの材料，中間成果物，成果物が高い価値を持ち，ユーザ・ベンダ共にそれらを独占／再利用したいという意向を強く持つ。

たとえばユーザが提供した生データを用いてベンダが学習済みモデルを生成するという典型的なケースを前提とすると，ユーザとベンダの成果物に関する意向は以下のように対立することになる。

・ユーザ

学習用データセット，学習済みモデルは自社のノウハウや機密が詰まった生データを用いて生成されたものであり，開発に際して委託料も支払っていることから，自社で独占したい。

・ベンダ

生データを用いて学習用データセットを生成する過程，学習済みモデルの生成過程いずれにおいても自社の高度のノウハウ及び多大な労力を用いていること，学習用データセットや学習済みモデルは再利用が可能であることから，委託を受けた本開発案件以外にも横展開したい。

そこで，この相反する２つの意向を調整する枠組みが必要となる。

3. 2　調整の枠組み

筆者は以下のような枠組みで調整すればよいのではないかと考えている。

１　材料・中間成果物・成果物について，何が知的財産権の対象となるのか・ならないのかを知っておく

※本文の複製、転載、改変、再配布を禁止します。

２　１についてデフォルトルール（＝法律上のルール）として誰がどのような権利を持っているかを知っておく
３　契約条項をどのようにして自社に有利にデザインするかを知っておく（「権利帰属」にこだわらず「利用条件」で「実」をとる）

AI開発における材料，中間成果物，成果物の権利や知財に関して整理する際には，「知的財産制度という法制度上のルール」と「当事者間の契約による調整」という２つの視点が重要である。

ここで，AI開発におけるこの２つの視点の関係を整理すると以下の表３のようになる。

表３　「知的財産制度という法制度上のルール」と「当事者間の契約による調整」

法律の規定	当事者間の契約なし	当事者間の契約あり
知的財産権の対象となるもの ＝法律上の規定あり	知財法（特許法や著作権法）の規定（デフォルトルール）がそのまま適用され、権利者や権利内容が確定する	法律の規定が当事者の契約どおり修正される
知的財産権の対象ではないもの ＝法律上の規定なし	ルールなし （原始状態）	当事者の契約どおりとなる

まず「１　材料・中間成果物・成果物について，何が知的財産権の対象となるのか・ならないのかを知っておく」という点である。

これは，各成果物等が知的財産権の対象となるかならないかで，その帰属や利用条件についてデフォルトルール（＝法律上のルール）があるかないかが異なるためである。

次に「２　１についてデフォルトルール（＝法律上のルール）として誰がどのような権利を持っているかを知っておく」である。

つまり，まず「１」で何が知的財産権の対象になるかを知ったうえで，「２」で対象となるものについての法制度（特許法，著作権法など）上のルールを知っておくということである。

最も重要なのは，その次の「３　契約条項をどのようにして自社に有利にデザインするかを知っておく（「権利帰属」にこだわらず「利用条件」で「実」をとる）」である。

１，２はいわば「知っておくべき知識」であるが，３についてはまさに自社のビジネス上，成果物の権利帰属の要否やどのような利用条件が必要なのかを十分検討し，その検討結果を反映させた契約条項にする必要があるからである。

詳細は後述するが，ここでのデザインの仕方が，AI開発契約における「権利と知財」に関する考え方の肝であり，AIガイドラインにおける非常に大きなポイントの１つである。

3. 3　権利の所在を知っておく

「１　材料・中間成果物・成果物について，何が知的財産権の対象となるのか・ならないのかを知っておく」「２　１についてデフォルトルール（＝法律上のルール）として誰がどのような権利を持っているかを知っておく」の，２点はまとめて説明したほうがわかりやすいのでまとめる。

ここで検討する必要がある対象物（材料・中間成果物・成果物）は，以下の６つである。

１　生データ
２　学習用データセット
３　学習用プログラム
４　学習済みモデル
５　学習済みパラメータ
６　ノウハウ

これら６つの対象物を現行の知財法で保護しようとすると，実際には特許法，著作権法，不正競争防止法（営業秘密等）の３つが考えられる。

(1) 生データ

１）知的財産権の対象となるのか・ならないのか

※本文の複製、転載、改変、再配布を禁止します。

生データの種類によるが，例えば機械の操業データ，センサデータや事実を示すデータなどについて知的財産権は発生しないので，「営業秘密」（不競法２条６項）「限定提供データ」（改正不競法２条７項）（以下「営業秘密等」という）に該当する限りにおいて，不正取得等から保護されることになる。

営業秘密等にも該当しない生データについては，法律上のデフォルトルールがないということになる。

２）デフォルトルール（＝法律上のルール）として誰がどのような権利を持っているか

営業秘密等に該当しない生データについては，知的財産権の対象ではないため誰も権利を持っていない。したがって，そのような場合において生データを誰がどのように利用出来るかについては，ユーザ・ベンダ双方の契約によって定めるしかないことになる。

(2) 学習用データセット

「学習用データセット」とは，生データに変換・加工処理を施し，学習作業を容易にするために生成された二次的な加工データのことを言う。

１）知的財産権の対象となるのか・ならないのか

学習用データセットは，情報の単なる提示に過ぎないため，「発明」に該当せず特許を受ける権利の対象とはならないことがほとんどだと思われる。

一方，個々のデータに著作物性がない場合でも，学習用データセットが「データベースの著作物」（著作権法12条の２）に該当すれば著作権が発生する。

「データベースの著作物」とは「その情報の選択又は体系的な構成によって創作性を有するもの」を言うが，効率的な機械学習・深層学習のために，生データを取捨選択したり，体系的な構成で整理したりした学習用データセットは「データベースの著作物」に該当する場合もあると思われる。

また，営業秘密等に該当すれば不正競争防止法で保護される。

２）デフォルトルール（＝法律上のルール）として誰がどのような権利を持っているか

学習用データセットが「データベースの著作物」に該当する場合，創作的な「情報の選択」又は「体系的な構成」を行った者が著作権者となる。

したがって，ベンダのノウハウのみを利用して加工行為を行ったのであればベンダが著作権者となるし，ユーザとベンダが共同して創作的な行為を行ったのであればユーザ・ベンダの共同著作物となって双方が著作権を共有するということもありえる。

(3) 学習用プログラム

「学習用プログラム」とは，学習用データセットを利用して学習を行い，学習済みモデルを生成するためのプログラムを言う。

学習用プログラムは，ベンダが既に保有しているものを利用する場合や，具体的開発案件に即して一から開発する場合など様々なケースがあるが，実際には，OSS（オープン・ソース・ソフトウェア）が利用されることも多々ある。

１）知的財産権の対象となるのか・ならないのか

学習用プログラムは「プログラム」なので，通常のプログラムが知的財産権の対象となるかどうかと同義である。つまり，アルゴリズムは，特許法上の要件を充足すれば「物（プログラム）の発明」等として特許を受ける権利が発生するし，ソースコードは著作権法による「プログラムの著作物」として著作権法上の保護を受ける（なお，オブジェクトコードに変換されても同様。著作権法10条１項９号）

また，営業秘密等に該当すれば不正競争防止法で保護される。

※本文の複製、転載、改変、再配布を禁止します。

２）デフォルトルール（＝法律上のルール）として誰がどのような権利を持っているか

法律上，特許を受ける権利を取得するのは発明者（作成者）であり，著作権を取得するのは創作者（作成者）であるから，当該プログラムを発明・創作した者が特許を受ける権利・著作権を取得することになる。したがって，学習用プログラムをベンダが一から開発したのであれば法律のデフォルトルールとしてはベンダが，特許を受ける権利も著作権も取得することになる。

(4) 学習済みモデル

学習済みモデルは，学習用データセット同様，再利用可能であり，契約当事者の関心が非常に高い成果物である。

ただし，学習済みモデルに関しては，契約上・交渉上『学習済みモデル』という言葉が何を意味しているのかについて，慎重に見極める必要がある。というのは，学習済みモデルは，「関数」「数理モデル」「アルゴリズム」「ネットワーク構造」「推論プログラム」「パラメータ」「それら各概念の組み合わせ」等多義的な意味を持っており，当事者が異なる意味で使うと大きなトラブルの原因となるからである。本稿では，AIガイドラインと同様，学習済みモデルとは「『学習済みパラメータ』が組み込まれた『推論プログラム』」を指すものとしている。

１）知的財産権の対象となるのか・ならないのか

学習済みモデルのうち「推論プログラム」部分については，学習用プログラムと同様に考えれば問題ない。つまり，アルゴリズムは，特許法上の要件を充足すれば「物（プログラム）の発明」等として特許を受ける権利が発生するし，ソースコードは著作権法による「プログラムの著作物」として著作権法上の保護を受ける。たとえば，特定の開発課題の関係で非常に高い精度を持つ独自性の高いネットワーク構造を発見した場合，そのネットワーク構造に関するアルゴリズムについては「物（プログラム）の発明」として特許出願が可能となる可能性がある。

また，営業秘密等に該当すれば不正競争防止法で保護される。

学習済みモデルのうち「学習済みパラメータ」部分については後述するが，結論的には知的財産権の対象にはならないものと考える。

２）デフォルトルール（＝法律上のルール）として誰がどのような権利を持っているか

これも，「推論プログラム」部分については，学習用プログラムと同様，ベンダが一から開発したのであれば法律のデフォルトルールとしてはベンダが特許を受ける権利も著作権も取得することになる。

(5) 学習済みパラメータ

「学習済みパラメータ」とは，学習用データセットと学習用プログラムを用いた学習の結果，得られたパラメータ（係数）をいう。すなわち「学習用プログラムで自動的に生成される」かつ「大量の数値の列」であり，ディープラーニングの場合で言うと，学習済みパラメータの中で主要なものとしては，各ノード間のリンクの重み付け等がこれに該当する。

１）知的財産権の対象となるのか・ならないのか

先ほど説明したように，学習済みパラメータは，「学習用プログラムで自動的に生成される」かつ「大量の数値の列」であって創作性がないことから「発明」にも「著作物」にも該当しないものと思われる。

もっとも営業秘密等に該当すれば不正競争防止法で保護される。

２）デフォルトルール（＝法律上のルール）として誰がどのような権利を持っているか

営業秘密等にも該当しない学習済みパラメータについては，知的財産権の対象ではないため誰も権利を持っていない。したがって，学習済みパラメータを誰がどのように利用出来るかに

※本文の複製、転載、改変、再配布を禁止します。

ついては，ユーザ・ベンダ双方の契約によって定めるしかないことになる。

(6) ノウハウ

AIソフトウェアの開発に際しては様々なノウハウが生じる。たとえば，「実環境から生データを取得・選択する方法」「学習用データセットへの加工方法」「学習用プログラムを用いた効率的な学習方法」「学習済みモデルの本番環境における調整」などに関するノウハウである。

1）知的財産権の対象となるのか・ならないのか

ノウハウについては，無形の情報なので，著作権の対象にはならないが，「発明」の要件を満たすノウハウであれば特許を受ける権利の対象になりえる。また営業秘密等に該当すれば不正競争防止法で保護される。

2）デフォルトルール（＝法律上のルール）として誰がどのような権利を持っているか

営業秘密等や発明に該当しないノウハウについては，知的財産権の対象ではないため誰も権利を持っていない。したがって，ノウハウを誰がどのように利用出来るかについては，ユーザ・ベンダ双方の契約によって定めるしかないことになる。

(7) まとめ

以上をまとめると以下の表４のように整理できる。

表４　保護対象のまとめ

	特許法	著作権法	不正競争防止法
生データ	x	△ (著作物性があるデータのみ)	○ (営業秘密等の要件を満たす場合。以下同様)
学習用データセット	x	△ データベースの著作物に該当する場合	○
学習用プログラム	○	○	○
学習済みモデル	○ ただし推論プログラムのアルゴリズム	○ ただし推論プログラムのコード	○
学習済みパラメータ	x	x	○
ノウハウ	△ 「発明」の要件を満たす場合	x	○

3. 4　契約条項をどのようにして自社に有利にデザインするかを知っておく

これで，AI開発における６つの中間成果物・成果物についての法律上のデフォルトルールが分かった。

次に重要なのは，そのデフォルトルールを前提として，契約条項をどのようにデザインするかである。

(1) 典型的な暗礁乗り上げパターン

AI開発契約における，権利・知財に関する典型的な暗礁乗り上げパターンは以下のようなものである。

ユーザとしては，学習用データセットや学習済みモデルは，ユーザのノウハウや機密が詰まった生データを用いて生成されたもので，開発に際して委託料も支払っていることから，自らの権利を主張する。それに対してベンダも，生データだけでは学習済みモデルは生成できず，高性能なモデルができるのは，データの前処理やモデルの訓練過程いずれにおいてもベンダの高度なノウハウと多大な労力が不可欠であるとして，ユーザ及びベンダの双方が自らの権利を

※本文の複製、転載、改変、再配布を禁止します。

主張するパターンである。

(2) 整理の枠組み

このような対立は，ユーザ・ベンダいずれもが「成果物等は自社のものである」，言い換えると「成果物の権利を自己に帰属させる」ことに双方が固執することに主として起因している。

そして，このように「どちらが権利を持っているか」（権利の帰属）に双方がこだわっている限り永久に双方の溝は埋まらず，交渉に多大な労力と時間がかかり結局双方が競争力を失うことになる。

そこで，提案しているのが「権利帰属」と「利用条件」を分離して柔軟な条件設定をすることである。

たとえば学習済みモデルにつき，「１　ベンダに権利を帰属させた上で（「権利帰属」），「２開発後，ベンダは一定期間の目的外利用や競業的利用は禁止される一方でユーザは当該学習済みモデルを自由に利用できる（「利用条件」）」等の対応をすることによって，当事者双方の利益に合致する契約を締結できる場合もあるだろう。

言い換えれば「双方が対象物の『権利帰属』ではなく『利用条件』で『実』をとることを目指す」という発想である。

極端な言い方をすれば，自社が学習済みモデルに関する権利を保有しておらず，相手に権利が全て帰属していても，交渉の結果「モデルの第三者提供を含め，何の制限もなくモデルを自由に利用できる」という利用条件を設定できれば，実質的にはモデルの権利を保有していることとほとんど同じ，ということである。

(3) 具体的な検討方法

このように「権利帰属」と「利用条件」を分けて考えるという発想に立つと，理屈としては，６つの中間成果物・成果物全てについて「権利帰属」と「利用条件」を設定するということになる（なお，以下の図で生データについて「権利帰属」を定めていないのは，生データについては現行法上知的財産権が発生しないため，直接「利用条件」を定めれば足りるためである（もちろん，著作物など知的財産権が発生する生データについては「権利帰属」が問題となる））。

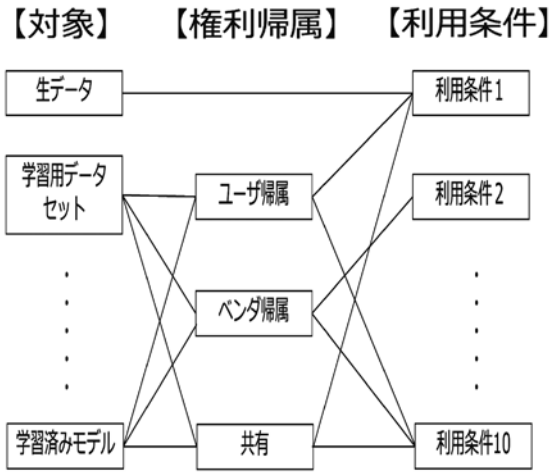

図４　中間成果物・成果物全てに権利帰属と利用条件を定めた概念図

もっとも，実際にはもっとシンプルなパターンも多いはずで，AIガイドラインに附属している「開発段階のソフトウェア開発契約書（モデル開発契約書）」では，シンプルな契約条項も提案している。

１）権利帰属について

権利帰属については，誰に権利帰属するかを合意するとすれば以下の３パターンしかない。

・ベンダ全部帰属
・ユーザ全部帰属
・ユーザ・ベンダ共有

である。

AIガイドラインのモデル開発契約書においては，成果物等のうち著作権の対象となるものは第16条に，著作権以外の知的財産権の対象となるものは第17条に定めている。

２）利用条件について

※本文の複製、転載、改変、再配布を禁止します。

利用条件については，ユーザ，ベンダそれぞれが，材料・中間成果物・成果物を，自社のビジネスにおいてどのように利用したいかをよく検討しなければならない。

たとえば，学習済みモデルの利用条件であれば，ユーザ・ベンダそれぞれがどのように自己のビジネスに利用するかは様々な方法が考えられるため，以下のような要素を考慮する必要がある。

・自己の業務遂行に必要な範囲で，開発対象となった学習済みモデルを利用するだけなのか
・学習済みモデルに新たなデータによる学習を行い，派生モデル（AIガイドラインでは「再利用モデル」としている）を生成するのか
・学習済みモデルや派生モデルを第三者へ開示，利用許諾，提供等することがあるのか
・相手方に対する利益配分（ライセンスフィー，プロフィットシェア）が必要かどうか

筆者としては実際の契約締結交渉においては「権利帰属」より，ここの「利用条件」をいかに自分のビジネスモデルに適合した形で設定できるか，ということの方が重要ではないかと感じている。

3. 5 小　括

以上述べてきたように「生成された学習用データセット，学習済みモデル，学習済みパラメータは誰がどのような権利を持っているのか（権利・知財）」という点については，AI開発契約の当事者が非常に強い関心を持つ領域である。それが故に双方の意向が対立し交渉に非常に長い時間を要することも多い。

その点を整理するための枠組みとして本稿では以下の３点を提案した。

(1) 材料・中間成果物・成果物について，何が知的財産権の対象となるのか・ならないのかを知っておく
(2) (1) についてデフォルトルール（＝法律上のルール）として誰がどのような権利を持っているかを知っておく
(3) 契約条項をどのようにして自社に有利にデザインするかを知っておく（「権利帰属」にこだわらず「利用条件」で「実」をとる）

4． 責任について

4. 1 ベンダがユーザに対して負担する３種類の責任

AIの開発・利用に関してベンダがユーザに対して負担する可能性がある責任は以下の３つに分類できる。

1　AI開発遂行に際してユーザに生じた損害についての責任
2　AIソフトウェアの利用によりユーザに生じた損害についての責任
3　AIソフトウェアをユーザが利用したことによりユーザが第三者の知的財産権を侵害した場合の責任

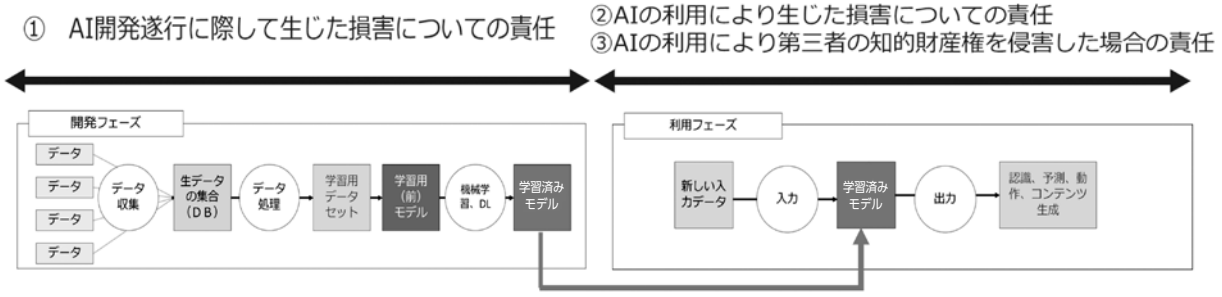

図５　AIの開発・利用の各フェーズにおける責任との関係図

※本文の複製、転載、改変、再配布を禁止します。

4. 2 AI開発遂行に際してユーザに生じた損害についての責任

AI開発遂行に際してユーザに生じた損害の具体例として，「ベンダが，通常の技術レベルを持つAIベンダであれば発生しないレベルのミスを犯したことにより，学習に通常では考えられない期間を要したため納期に間に合わなかった。」場合の責任を考える。

当然のことだが，AI開発契約の法的性質を準委任契約としたからといって，受任者であるベンダが一切責任を負わないということではない。準委任契約においても，ベンダは「善管注意義務」（民法644条。委任の本旨に従い，善良な管理者の注意をもって，委任事務を処理する義務）を負っている。したがって，この「開発遂行に際して生じた責任」についてはAI開発と通常のシステム開発を区別する合理性がない。

そこで，AIガイドラインに添付されているモデル開発契約第22条１項では以下のように定めている。

【第22条１項】

「ユーザおよびベンダは，本契約の履行に関し，相手方の責めに帰すべき事由により損害を被った場合，相手方に対して，損害賠償（ただし直接かつ現実に生じた通常の損害に限る。）を請求することができる。ただし，この請求は，業務の終了確認日から●か月が経過した後は行うことができない。」

このうち「ただし直接かつ現実に生じた通常の損害に限る」という部分については，民法に定められている損害賠償の範囲（民法416条）を制限するものでありベンダにとって有利な規定となる。そのため，委託料の金額や発生する可能性のある損害の性質・金額によっては，そのような制限を外し，民法の原則に戻すことが合理的となることもある。

4. 3 AIソフトウェアの利用によりユーザに生じた損害についての責任

成果物であるAIソフトウェアの利用によりユーザに生じた損害の具体例として，「工場における半製品の異常検知AIをベンダが開発してユーザに納品，ユーザが自社の工場において当該AIを利用したところ，AIソフトウェアが異常を見落としてユーザが不良品を顧客に出荷してしまい大きな損害を被った。」場合の責任を考える。

先ほどの「AI開発遂行に際してユーザに生じた損害についての責任」と異なり，「AIソフトウェアの利用によりユーザに生じた損害についての責任」については，ベンダにその責任を問うことがかなり難しいのではないかと思われる。

これは「因果関係等につき事後的な検証等が技術上困難である」「学習済みモデルの性能等が学習用データセットに依存する」「AIソフトウェアの出力が利用段階の入力データの品質に依存する」という点に由来するものである。

したがって，AI開発契約においては，この「AIソフトウェアの利用によりユーザに生じた損害についての責任」については，ベンダは責任を負わないとするのが合理的ではないかと思われる。

そこで，モデル開発契約第20条では，学習済みモデルなど成果物等の使用等による責任をベンダは原則として負わないと定めている。

【第20条】

「ユーザによる本件成果物等の使用，複製および改変，並びに当該，複製および改変等により生じた生成物の使用（以下「本件成果物等の使用等」という。）は，ユーザの負担と責任により行われるものとする。ベンダはユーザに対して，本契約で別段の定めがある場合またはベンダの責に帰すべき事由がある場合を除いて，ユーザによる本件成果物等の使用等によりユー

※本文の複製、転載、改変、再配布を禁止します。

ザに生じた損害を賠償する責任を負わない。」

もちろん，当事者のニーズや力関係によっては，学習済みモデルの利用によって生じた損害についてもベンダに責任を負って欲しいということはあり得ると思われる。

そのような場合においては，一定の期間や上限金額を設けた上でベンダが損害賠償責任を負担するという条項にすることも考えられる。また「学習済みモデルの利用によって生じた損害」のうち，第三者の知的財産権を侵害したことによって生じた損害についてはベンダが責任を負うことが合理的な場合もあるので，別途の考慮が必要となる（詳細は後述）。

4. 4 AIソフトウェアの利用によりユーザが第三者の知的財産権を侵害した場合の責任

成果物であるAIソフトウェアをユーザが利用したことによりユーザが第三者の知的財産権を侵害した場合の具体例として，「ある学習済みモデルの生成方法（学習方法）についてA社により特許登録がなされていた。ベンダが当該特許発明をA社に無許諾で実施して学習済みモデルを生成してユーザに提供し，ユーザが不特定多数の第三者に当該モデルを提供し始めたところ，A社からユーザに特許権侵害であるとの警告書が届いた。」場合の責任について考える。

この「AIソフトウェアの利用によりユーザが第三者の知的財産権を侵害した場合の責任」は「AIソフトウェアの利用によりユーザに生じた損害」の一種だが，知的財産権侵害についてはユーザの関心が非常に高いため別に検討する必要がある。

ユーザとしては，ベンダに対して，知財の非侵害保証（第三者の知的財産権を侵害しないことの保証）をしてほしいという要請を行うこともあるが，一般論としては，ベンダにおいて海外特許を含め侵害の有無を完全に調査検証することは費用面から非常に困難であることも少なくない。

そこで，大まかな考え方としては，以下の３つがありえる。

１　一切保証をしないパターン
２　著作権非侵害のみ保証するパターン
３　すべての知的財産権の非侵害を保証するパターン

モデル開発契約においては，上記２（第21条Ｂ案）と３（第21条Ａ案）を提案している。AI開発に用いられている手法や，当該手法の特許権侵害の可能性等を考慮し，上記１～３のいずれかを選択することとなろう。

5. おわりに

以上，AIソフトウェア開発契約締結に際して問題となることが多い３つの領域，具体的には「AI開発契約において「性能保証」「検収」「瑕疵担保」についてはどのように定めればいいのか（性能保証，検収，瑕疵担保）」「生成された学習用データセット，学習済みモデル，学習済みパラメータは誰がどのような権利を持っているのか（権利・知財）」「AI開発・利用に際して生じる可能性のある損害についてAI開発契約ではどのように定めたらよいか（責任）」について，その内容と契約上の対応方法についてAIガイドラインをベースに解説を行った。

AIソフトウェアの開発においては，通常のシステム開発と原理的に異なる点が多々あることから，契約当事者双方がこれまでの契約交渉セオリーをそのまま振りかざしたのでは合理的な交渉とならないことが多い。

契約交渉に際しては，本稿で紹介したような視点から，自社のビジネス推進のために，交渉条件のどこを譲ってどこを堅持するかについての徹底した検討が非常に重要となる。

※本文の複製、転載、改変、再配布を禁止します。

注　記

1）AI・データの利用に関する契約ガイドライン
http://www.meti.go.jp/press/2018/06/20180615001/20180615001.html
（URL参照日：2019年３月７日）

2）なお，本記事はAIガイドラインを題材にした筆者個人の見解であって，「ガイドラインの概要」「ガイドラインのエッセンス」「ガイドラインの公的な解釈」ではないことを，あらかじめお断りする。

3）「PoC」とは「Proof Of Concept」（概念実証）のことである。一般的には，新しい概念やアイデア，手法などを，実用化できるかどうかを検証することをいうが，AI開発においては学習用データセットを生成し，ユーザが希望する精度の学習済みモデルが生成できるかを検証することを意味する。

4）成果完成型といってもあくまで準委任契約なので，一定の成果（例：精度）を出すことが契約上の義務となるわけでない。したがって，成果完成型の準委任契約を締結した場合において，成果が達成できない場合でも，ベンダが報酬を受領できないだけであって，請負契約と異なりユーザに対して債務不履行責任などの法的責任を負うわけではない。

5）この図３における「成果物」とは契約上，納品や作成支援が合意されているものを指す。したがって「成果物」と「中間成果物」の区別は相対的なものであり，契約内容によっては学習用データセットが成果物として合意されることもある。

（原稿受領日　2018年12月17日）