AI 学習履歴管理システムの提案
～ブロックチェーンを使った AI 血統検証基盤～

川本　康貴†　小林　啓洋†

†沖電気工業株式会社 経営基盤本部 研究開発センター 〒542-0051 大阪市中央区備後町 2-6-8
E-mail:　†{kawamoto728,kobayashi050}@oki.com

あらまし　AI が社会に受け入れられて社会基盤技術となるためには，AI の技術的側面に加え，AI の倫理的側面も重要になる．本稿は，「手元の AI が，どのような AI を元にどのようなデータを使って学習をしたものか」を検証するための，AI 血統検証基盤の提案である．このようなシステムを既存のデータベースシステムを使って構築すると，システムの可用性や登録内容の完全性の観点から問題がある．そこで本提案のシステムはブロックチェーン基盤である Hyperledger Fabric を使って構築した．このことにより，本提案のシステムはシステムの可用性や登録内容の完全性を確保できる．ブロックチェーン内のデータは完全性が担保できるので，本システムを利用すればオープンデータを使った自動 AI モデル生成や AI モデルの自動売買ができる．

キーワード　AI 倫理, ブロックチェーン

Proposal of the AI learning history management system
-　AI pedigree verification platform using blochchain-

Yasutaka KAWAMOTO†　　Akihiro KOBAYASHI†

†Corporate Infrastructure Group, Corporate Research & Development Center, Oki Electric Industry Co., Ltd.

2-6-8 Bingomachi,Chuo-ku,Osaka 541-0051,Japan
E-mail:　†{kawamoto728,kobayashi050}@oki.com

Abstract In order for the AI to be accepted by society and become a social infrastructure technology, the ethical aspect of AI becomes important. The ethical aspects of AI include accountability for what data the AI has learned. This proposal is a proposal for the AI pedigree verification platform, which is a system for verifying "what the AI at hand has learned". If such a system is constructed as an existing database system, there are problems from the viewpoint of ensuring system availability and ensuring the integrity of registered contents. Therefore, the proposed system was constructed using Hyperledger Fabric, which is a blockchain platform. By using blockchain, the proposed system can ensure system availability and integrity of registered contents. Data stored in the blockchain can be guaranteed integrity. Therefore, we can realize automatic generation of AI models using open data and automatic trading of AI models by using the system.

.

Keyword　AI ethics,　blockchain

1. はじめに

現在，我が国では少子高齢化による労働人口の低下が懸念されている．例えば，人口問題研究所の資料[1]によれば，2015 年には 7000 万人程度(全体の 60%)だった日本の労働者人口(15 歳～65 歳)が，2050 年には 5000 万人(全体の 50%程度)にまで低下する．労働人口が減ると国単位でみた生産性が落ち，国力が低下する．そこで，少ない労働人口でも高い生産性や付加価値を生み出すために，ICT（Information and Communication Technology）の活用が期待されている．

内閣府の発表[2]によれば現在は，IoT（Internet of Things）技術で集めた実社会（フィジカル空間）の情報を，サイバー空間のクラウドサービスに集め，人がその情報を分析して，社会へフィードバックするという Society 4.0 である．今後は，IoT 技術で集めたフィジカル空間の情報をサイバー空間で統合して AI（Artificial Intelligence）を使って自動解析して結果を社会にフィードバックすることで，自動的に新たな価値を創造する「Society 5.0」になると考えられている．

Society4.0 と Society5.0 の大きな違いは，集めた情報を人ではなく AI に解析させる点である．そのため，AI が社会的に受け入れられていないと，AI の解析結

果を価値に変換できず Society4.0 から 5.0 へのパラダイムシフトが起こらない．このことから，AI 作成者は AI の判定精度や学習速度といった技術的な側面だけではなく，AI が社会に与える影響や AI の判定結果に関する説明責任といった，AI の倫理的な側面（以下，AI 倫理）に関しても責任を負う必要があると言われている．例えば文献[3]では，AI が社会に受け入れられるためには，「AI に利用されるデータの取得方法や使用方法，AI の動作結果の適切性を担保する仕組み」が必要と記載されている．また，AI モデルを組み込んだ製品を販売しようとしているメーカも AI 倫理に関する独自の基準を設けている．例えば文献[4]は AI 事業に関して「説明と透明性」「対話と協調」といった AI の倫理的側面に注力することを宣言している．

本稿は，ある AI モデルが，どのような AI モデルを元にどのように学習を経て作られたものかを，AI 作成者が透明性をもって説明するための「AI 血統検証基盤」の提案である．

AI 作成者は，AI モデルを作る際の元となった AI モデルや学習に使ったデータ，学習アルゴリズムといった AI モデルの学習に関する情報（以下，AI 学習履歴情報）を本システムへ登録しておくことで，自身が作った AI モデルの AI 学習履歴情報，つまり AI の血統に関する説明責任を担保できる．また AI 利用者は本システムを使って手元にある AI モデルの AI 学習履歴情報を検証できるので，AI モデルを安心して利用できる．

AI 血統検証基盤は検証結果の公平性，透明性の確保が重要になる．また，今後，社会の基盤技術となる AI の血統を明らかにするという重要な機能を担うため，システムとしての可用性も確保する必要がある．

しかし，本システムを従来技術で構築した場合，AI 学習履歴情報を登録者は，一旦システムに登録した AI 学習履歴情報を後から偽造できる．これでは AI 血統の検証結果を信用できない．また，従来技術でシステムを構築した場合，システム運用者が運用をやめるとシステムが使えなくなる．これは社会基盤である AI を検証するシステムの可用性としては問題がある．

そこで，本システムはブロックチェーンを利用して構築した．具体的にはブロックチェーン開発基盤の一つである Hyperledger Fabric[5]を利用した．例えば文献[6]ではブロックチェーンの特性として『改ざんが極めて困難』を挙げている．この特性により，本システムの学習履歴情報の偽装は困難であり，検定結果の公平性や透明性を確保できる．また，Hyperledger Fabric は様々な会社が共同で運用する「コンソーシアム型システム運用」が可能なので，システムの可用性も確保できる．

2. 関連技術と課題

2.1. ブロックチェーン

ブロックチェーンとは，ハッシュ関数と電子署名を利用した分散型電子台帳技術である[7].

ブロックチェーンのデータは，あらかじめ決められたブロック単位で保存される．各ブロックは一つ前のブロックのハッシュ値を内包している．あるブロックのデータを改ざんすると，そのブロックのハッシュ値は変わる．各ブロックはハッシュ値で連結されているため，あるブロック内のデータを改ざんしようとするとそれ以降のブロックのハッシュ値がすべて変わってしまう．これは，あるデータを改ざんするためにはそれ以降のブロックをすべて改ざんする必要があるということである．このようなデータ構造をチェーン構造という．ブロックチェーンはチェーン構造により，内部データの完全性（Integrity）を保証している．

ブロックチェーンは，ブロック追加時に PoW（Proof of Work）[8]や PBFT（Practical Byzantine Fault Tolerance）[9]といったコンセンサスアルゴリズムを使って，既存のブロックへの不正なブロックの追加（二重投稿等）を防いでいる．例えば Bitcoin[8]が採用している PoW では，ブロック追加時に，ブロックと nonce(number used once)を加えた値のハッシュ値が，システムの指定する値以下である nonce を発見するまでブロックの追加ができない．こういった nonce を探す処理を Bitcoin では「マイニング」と言われており，マイニングを実施する計算機を「マイナー」と呼ぶ．マイニングが完了するとブロックの追加ができる．マイニングには多くの計算処理が必要なので，多数のマイナーが協力してマイニングをする．一部の不正なマイナーが結託して不正なブロックを追加しようとしても，Bitcoin を運用するためにマイニングを実施している多数の正当なマイナーに比べてマイニングに時間がかかる．その結果，不正なブロックの追加は，正当なブロックの追加よりも時間がかかってしまうので，システムとして「先にマイニングが完了したブロックを正しいブロックとみなす」と定義すると，結果として不正なブロックの追加は実施されない．

コンセンサスアルゴリズムとして PoW 以外によく利用されているものとして Hyperledger Fabric で採用されている PBFT がある．PBFT では，追加したいブロックに対して不正がないかどうかチェックし，不正がない場合にブロックに対して信頼できる計算機の電子署名をすることによって，ブロックの追加を許可する方式である．コンセンサスアルゴリズムの観点からすると，PBFT は PoW に比べてシステム内の計算処理能力が少なくても高速に運用できるところが強みである．逆に PBFT は，ブロック追加のために信頼できる計算

機が必要なので，Bitcoin のようなパブリック型のシステムには向かない．

コンセンサスアルゴリズムとしてどちらの方式を採用する場合でも，ブロックチェーンでは複数の計算機が協力してアルゴリズムを実行している．これをシステム運用の観点からみると，ブロックの追加というシステムの重要な部分を複数の計算機で分担することで冗長性をもって運用していることになる．このことから，ブロックチェーンシステムは一部の計算機がシステム運用をやめてもシステムの運用は継続できるので，可用性(Availability)が担保できるといえる．

ブロックチェーンは，コンセンサスアルゴリズムで構築したチェーン構造のデータを，P2P（Peer to Peer）ネットワークを利用して複数の計算機に分散保存することで，システム内部のデータの信頼性（Reliability）を担保している．

2.2. 競合技術：DVC（Data Science Version Control System）[10]

本提案と似た技術として DVC がある．

DVC は AI モデルを生成する際の学習環境の構成管理システムである．

一般的に，AI モデルは AI モデルの学習方法等を記述したプログラムコードと，学習に利用する学習データを AI モデル生成基盤に入力することで得られる．

DVC でいう「学習環境」とは，AI モデル生成のためのプログラムコードと学習データを指す．このうち，プログラムコードは従来の構成管理システム[11]を利用して保存・管理できる．しかし，学習データは巨大（数十ギガバイト以上）であることが多いので，従来の構成管理システムに保存することは馴染まない．そこで DVC では学習データ保存用のストレージを別途用意して学習データを保存し，学習データの保存場所と AI モデルの構成を記述したファイル（dvc ファイル）作成する．dvc ファイルはプログラムコードと統合して従来の構成管理システムで管理する．学習環境を再現したい場合，まず，従来の構成管理システムからプログラムコードと dvc ファイルを取り出し，dvc ファイル内部の情報を使って関係する学習データを取り出す．

DVC を使えば AI モデルがどのような学習環境で生成されたかの管理はできる．しかし，DVC を AI モデル作成者が AI モデル利用者に対して，AI モデルをどのように学習させたかを説明するための「AI 血統検証基盤」とした観点でとらえた場合，以下のような課題がある．

●AI モデルからの学習環境の逆引き機能がない

DVC は AI モデル作成者が自身の学習環境を管理するためのツールである．しかし，AI モデル生成のために利用したプログラムコードや学習データを検証したいのは AI モデル作成者だけではない．例えば AI モデル利用者の観点からすると，利用中の AI モデルがどのような学習データやプログラムコードで生成されたものか知りたいという要求はある．DVC はこういった要求には対応できない．

●システムの可用性確保が困難

DVC の構成管理システムとして，Git[11]等の分散型構成管理システムを採用し運用することで，分散化によるデータの信頼性は担保できる．しかし，長期間運用されるという基盤としての可用性は担保できない．なぜなら分散型構成管理システムは，データの分散できるが，システムとしての構成は中央集権型だからである．例えば，Git を使ったソースコード共有基盤である GitHub[12]は，ソースコードは Git によって分散化しているのでデータの信頼性は担保している．しかし，GitHub をシステムの観点で見た場合，GitHub の運用者がサイトの運用をやめると，GitHub システムは利用できなくなる．今後，AI モデルが社会のインフラとして浸透していくと，AI 血統検証基盤は「AI モデルのインフラ」として AI モデル作成者や AI モデル利用者に利用される．よって，基盤の可用性が単一のサーバ運営者に依存していることは，システムの性格上望ましくない．

●登録内容の完全性保証が困難

既存の構成管理システムはソフトウェア開発者が利用することを想定している．よって，使い勝手を優先し，構成管理システムへの登録を取り消す機能が用意されている．例えば git の場合，リポジトリへの変更を取り消すコマンドとして，reset コマンドが用意されている．DVC は既存の構成管理システムを利用しているので，AI 作成者はシステムへいったん登録した AI モデル学習環境の情報を取り消すことができる．これは，AI 作成者が AI 学習環境の構成管理をするという観点からすると問題はない．しかし，AI 血統鑑定システムとして利用する場合，AI 作成者や AI 利用者の観点から問題がある．なぜなら，DVC に登録したプログラムコードや学習データが改ざんされていないことをシステム的に保証できないからである．このため，AI 作成者は DVC を使うことによって AI 利用者に対し，提供した AI モデルをどのようなプログラムコードと学習データで作成したかの説明責任は担保できない．

3. 提案システム

AI 血統検証基盤の機能は「AI 学習履歴の保存および開示」である．現時点での「AI 学習履歴情報」とは少なくとも，AI モデル，AI 学習データ，プログラムコードのハッシュ値と保存場所へのポインタを含む．

図 1 は提案する AI 血統検証基盤と利用者の関係を示したものである．現時点で想定している AI 血統検証基盤の主な利用者は，AI モデルを作成して本基盤に AI 出自情報を登録する「AI 作成者」と，他者の AI モデルの作り方（データの整理・利用方法等）を参考にしたい「AI 技術学習者」と，AI 作成者が作成した AI モデルを安心して利用したい「AI 利用者」である．

AI 作成者は作成した AI モデルの AI 出自情報を本基盤へ登録（①）することで，自身が学習させた学習済み AI モデルの AI 出自情報を管理できる．AI 技術学習者は本基盤から AI 出自情報を読み出す（②）ことで，AI 作成者の AI 作成方法を参考にしながら学習ができる．AI 利用者は AI 作成者から提供（③）された AI モデルのハッシュ値を計算し，本システムが開示する AI 学習履歴と比較する（④）ことで，自身が利用している AI モデルが想定しているものかを検証できる．

提案システムを AI 血統検証システムという観点から DVC と比較した場合，以下のような利点がある．

- ●AI モデルからの逆引きが可能
  提案システムは AI 学習履歴情報を，AI モデルのハッシュ値により検索できる．このことより，AI 利用者は独自に AI 学習履歴情報を参照し，利用している AI モデルの血統を検証できる．
- ●コンソーシアム型運営により，システムの可用性確保が可能
  提案システムはブロックチェーンを利用して構築しているため，コンセンサスアルゴリズムを複数社で分散して運用することができる（コンソーシアム型運用）．このことより，ある会社がシステム運用をやめてもシステムは他の会社が継続して運用できる．
- ●登録内容の完全性をシステムとして保証可能
  ブロックチェーンのデータ構造により，システムに記憶された情報は書き換えられていないことを保証できる．

4. 今後の展望

2.1 節で論じたとおり，ブロックチェーンはチェーン構造でデータを保存するので，内容が改ざんされていないことを機械的に検証できる．この事により，ブロックチェーン内に保存されているデータの自動流通や自動決済ができる．例えば文献[13]によれば「ブロックチェーンがインフラとして普及すれば，マシンによる自律的な経済活動が行われるようになる」旨が記載されている．以上のことから本提案の AI 血統検証基盤を使えば，オープンデータを使った AI モデルの自動生成や AI モデルの流通が可能である．

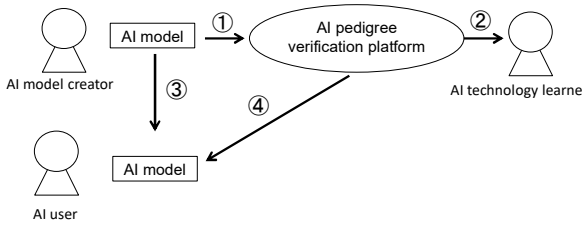

Fig.1. Relationship between the AI pedigree verification platform and users.

謝辞

本提案は大阪大学大学院情報科学研究科 教授 若宮直樹先生および助教 橋本匡史先生との議論中に着想を得，発展させたものです．お二人に感謝します．

文　献

[1] 人口問題研究所：
http://www.ipss.go.jp/pp-zenkoku/j/zenkoku2017/pp29_Report3.pdf
[2] Society5.0：
https://www8.cao.go.jp/cstp/society5_0/index.html
[3] 人間中心の AI 社会原則：
https://www8.cao.go.jp/cstp/aigensoku.pdf
[4] OKI グループ AI 原則：
https://www.oki.com/jp/press/2019/09/z19033.pdf
[5] Hyperledger Fablic:
https://www.hyperledger.org/projects/fabric
[6] IoT におけるブロックチェーンの適用可能性について:
https://www.jri.co.jp/MediaLibrary/file/report/jrireview/pdf/9954.pdf
[7] ANTONOPOULOS, Andreas M. Mastering Bitcoin: unlocking digital cryptocurrencies. " O'Reilly Media, Inc.", 2014.
[8] S. Nakamoto, Bitcoin: A peer-to-peer electronic cash system. 2008.
[9] M. Castro, B. Liskov, Practical Byzantine fault tolerance and proactive recovery. ACM Transactions on Computer Systems (TOCS), pp.398-461, ACM, 2002.
[10] DVC: https://dvc.org/
[11] git: https://git-scm.com/
[12] GitHub: https://github.com/
「ブロックチェーン技術の応用に関する戦略策定」報告書：http://www.glocom.ac.jp/news/3513

注：文献の URL は 2019/10/16 にアクセス確認をしました

注：Hyperledger は The Linux Foundation の商標です