

ディープラーニング開発エコシステム

橋 素子 前野 蔵人

ディープラーニングはAIに一大ブームをもたらした技術である。概念の登場から10年以上たった現在では、先進的な研究開発の発展と平行しコモディティ化も進み、開発の定石とも言える実装パターンの社会的蓄積が進んだ。これが、社会実装事例の増加と定石蓄積の循環を拡大し、PoC (Proof of Concept) ブームへと繋がっている。

ディープラーニングのモデルは、従来のAI技術に比べて汎用性が極めて高く、学習させるデータセットに合わせて実現する機能が柔軟に変化する。そのため、従来AI開発に必要であった専門家のもつ高い分析スキルが不要となり、質の良い大量のデータとソフトウェア開発スキルがあれば、誰もが高度なAIを開発できるようになった。

我々は、このパラダイム変化に合った開発を行える組織に変革する必要性を感じ、OKIグループ全体を対象に「ディープラーニング開発エコシステム」を提唱し活動している。事例の増加に呼応して増えるAI需要への対応と実績蓄積による認知拡大のサイクルを狙ったもので、教育啓蒙やコミュニティー、技術開発支援の活動を柱としている。本稿では、これらの活動について紹介する。

ディープラーニングによるパラダイム変化

ディープラーニングとはニューラルネットワークを多層化したもので、2012年に国際的な画像認識技術のコンペティションで認識精度を一気に10%も改善し注目を集めた。これをきっかけに研究が活発化し、その翌年以降コンテストの上位にランクインするものは全てディープラーニングに置き換わり、2015年には人間を上回る認識精度に達成した。その後もディープラーニングの社会実装は年々増加し、研究動向もブーム当初の画像認識からさまざまな技術領域に分化し、今なお拡大を続けている。

ディープラーニングの研究開発は世界的にオープンな文化の中で発達してきた。論文は学会発表前にプレプリントサーバーで公開され、そのソースコードは公開リポジトリで世界的に共有される。また、2015年を境に無料で簡単に使えるディープラーニング開発フレームワークやライブラリーが相次いで公開され、誰もが最新の技術に触れて実

証できる環境となった。近年では、異なるフレームワークで開発したディープラーニングモデルを標準的なフォーマットを経由して変換することが可能となり、モデルの流通が容易になったほか、試行と展開のそれぞれに適したフレームワークを自由度高く使い分けられるようになった。こうして開発の効率が劇的に改善し情報共有が進んだことで、社会的に検証の積み重なった実装パターンが定石として広く認知され、多様な用途での実装が進むようになった。

ディープラーニングが従来の機械学習技術と大きく異なる点は、大量のデータを学習することでAI処理に必要な特徴量を自動的に獲得できることにある。従来は、限られたデータを分析し有効な特徴を抽出するスキルが必要であったが、ディープラーニングでは、質の高いデータを大量に準備すれば、誰もが高度なAI処理を実現できる。開発環境の整備が進んだ現在では、Pythonなどのプログラミング言語を記述できるスキルがあれば、数学的知識や複雑な実装を必要とせず、定石となったレベルまでのディープラーニングを簡単に活用することができる。

これら開発スタイルの変化と同時に、学習に用いる大量データの質的マネージメントや、モデルとデータの関係性の管理が重要となってきている。また事例が増えるに従い、AIが下す判断の妥当性やサービス展開後の対応の重要性も頻繁に指摘されるようになった。こうしたパラダイムの変化は、リスクと見るだけでなく、日本企業が得意とする品質管理や保守の考え方も親和性が高く、今後の強みが発揮できる領域とも捉えられている。

ディープラーニング開発エコシステム

「ディープラーニング開発エコシステム」は、AI開発能力強化とその実績蓄積による認知度向上の循環を狙った活動である。AIのパラダイム変化に合わせたソフトウェア開発スタイルと自社の開発スタイルのギャップに着目し、それを埋めながら効果的に進められるように設計したもので、先駆的にディープラーニングを研究開発に取り入れてきたスペシャリストを中心に2017年に開始した。主要な活動は、図1に示すとおり(1)教育啓蒙、(2)情報共有コミュニティー、

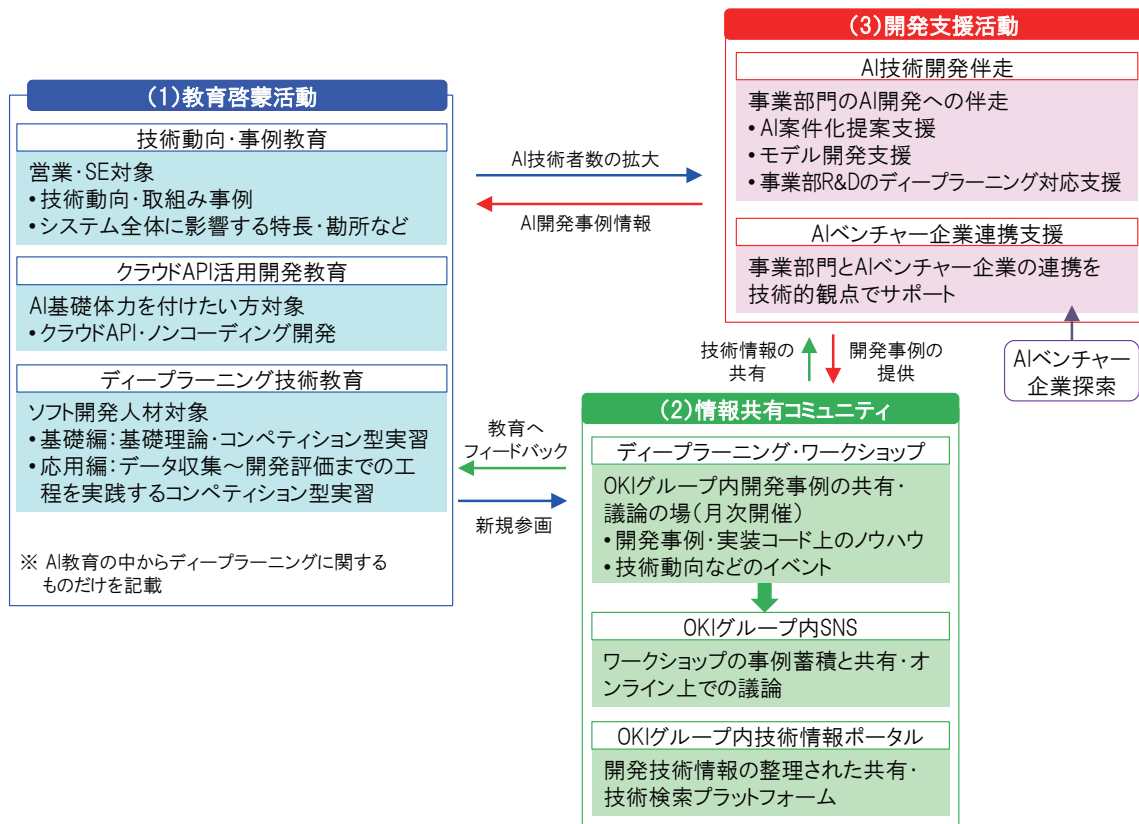


図1 ディープラーニング開発エコシステム概要図

(3)開発支援の3つである。(1)教育啓蒙では、座学によってディープラーニングの基礎知識を習得する。しかし、座学だけでは技術定着が難しい。このため、(3)開発支援により、事業部門の具体的な案件を支援し、技術者への確実な技術定着を促進する。こうした実経験は(2)情報共有コミュニティにより共有し、OKIグループ全体のAI技術レベルの向上を加速する。これら3つの活動を相互に深く関係性を持たせながら活動サイクルを回すことで、AI開発の定石を打てる人材を育成し、自社の開発スタイルをパラダイム変化に合わせたものへと変革している。以降では、それぞれの活動内容を詳しく説明する。

(1)教育啓蒙活動

ソフトウェア開発人材のAI人材化と営業・SEへのディープラーニング知識強化のための教育啓蒙活動を実施している。具体的には、ソフトウェア開発スキルを有する技術者へのディープラーニング技術教育と、営業やSEなどへの技術動向・事例教育である。

技術教育は、Pythonなどの開発言語の知識のあるソフトウェア開発者を前提とし、ディープラーニング初心者学ぶ基礎編とデータの準備から開発全体の工程を学ぶ応

用編を用意している。基礎編ではディープラーニングの基礎理論を学ぶと共に、その知識を活用してモデルを試作して精度を競うコンペティションを実施する。また応用編では、画像認識モデルの開発を題材に、画像の収集・正解ラベル付け・モデル設計・学習・推論・結果考察までの一連の開発の流れをグループ実習で実践する。基礎的な知識は与えつつ、認識精度で競わせるなど実践的な競争体験を通じて、技術者は技術開発プロセス全体を体得し、チューニングの勘所までを短期間で身につけることができる。

これら技術教育のカリキュラムは、実務でディープラーニングの研究開発を行っているメンバーが自身の経験から気づいた重要と感じる観点を入れているため、効率よい学びが得られる。実際に、これまでの受講者アンケートでは9割の参加者がその学習効率に高い評価を与えている。これまで開発スキルの高い技術者層からのべ130名以上が受講し、現在は受講者が自部門で講師となって講習会を行うなど二次的な広がりも見せている。

また技術教育では、上記よりも更に平易な入門編として、クラウド上で利用可能なAPI(Application Programming Interface)を活用するカリキュラムを事業部門主体で用意している。精度の良いモデルが部品化し外部から利用可能

となってきた現在の状況を踏まえ、これらの利活用知識の浸透もAI教育上必要な観点となっている。

営業・SE向けの技術動向・事例教育では、ディープラーニングの技術的特長や市場動向に加え、開発のパラダイムが変化したこと起因する事業上・開発上の留意点と開発事例などの解説を用意している。こちらでも技術に携わる中で営業部門やSE部門から得られるフィードバックを踏まえた工夫を入れ、既に300名以上が受講し、OKIがAIを事業に積極的に活かす風土醸成に貢献している。

(2) 情報共有コミュニティ

実践的な技術情報やノウハウを共有し、AI開発スキルを効率よく高める目的でコミュニティを構築している。具体的には、OKIグループ内研究会であるディープラーニングワークショップと、OKIグループのイントラネット内に設置されたSNSと技術情報ポータルサイトによる情報共有である。ディープラーニングワークショップを中心とした活動と位置づけ毎月定期的に開催し、多様な部門間でディープラーニングの開発体験を共有・議論する風土をつくり、恒常的に技術レベルとモチベーションを高めていくことを狙っている。

ここで報告される内容は幅広く、事業部門や研究開発部門の開発活動や先端の技術動向、開発環境やOKIグループ内の開発支援体制の議論などである。開発活動の状況では、具体的なコードレベルでの議論や開発プロセス上の苦労、落とし穴などの実体験に基づく話題が中心となり、机上では得られない生のノウハウ的な知見に触れられる。経験値の高い参加者との議論からは、多くの気づきも得られる。発表資料は随時SNSにアップロードされ、参加者はいつでも内容を確認し、時には疑問点を研究会だけではなくSNS上でディスカッションすることもできる。またこうして確立した技術情報は、技術情報ポータルサイトに整理・掲載され、多様な階層への技術情報の横展開が図られる。

このような直接的な事業貢献でないコミュニティ活動を長期継続することは難しく、試行錯誤を経て常に改善と工夫を繰り返しながら続けている。運営は筆者らが所属する研究開発センターが担い、最先端技術を研究開発している大学の先生を招いた技術セミナーや、AI知財戦略のディスカッションなどもワークショップの特別セッションとして実施している。これらは幅広い人材が常に好奇心をもって活動を続けられるための工夫の一つである。こうして、OKIグループ全体の技術力底上げを図っている。

(3) 開発支援活動

研究開発センターでは、事業部門が実施する具体的なAI関連事業に対し、顧客提案から試作・開発までの全体に

専門技術的視点で支援を実施している。どのようなデータを収集すれば良いか、今あるデータなら何がどのレベルまでできるか、データは膨大にあるがポイントとなる事象は少ないなど、多様な前提条件により取るべき最善のステップは提案段階から変化する。それらをディープラーニングのスペシャリストの見地で助言し、より効果的なアプローチに結び付けている。モデル開発では、基本は定石となっている手法をどのように適用していくかではあるが、実際の現場の多様な課題により臨機応変に対応している。例えば、前述のデータの状況によって採用すべきモデルの特長や実現すべき機能も変化し、さらには本格開発前のPoC対応の必要性なども考慮すべきポイントになる。こうした多様な開発体験を、事業部門の人材と伴走し共感することで、高い応用レベルでスキルを移転することができる。

この開発支援の枠組みの中では、必要に応じてAIベンチャー企業とも連携している。AI技術開発の全てを社内のリソースだけでまかなうのではなく、AI開発能力の高いベンチャー企業を活用しつつ、これまでのビジネスで培ったドメイン知識やデータへの知見を活かした開発でシナジーを作っていく観点も大切である。こうした考え方で、継続的に多数の企業を調査し、具体的な連携に結びつけている。例えば、創屋株式会社殿や株式会社Laboro.AI殿などと、複数のAIプロジェクトを連携して進める関係を構築している。

これまでの活動事例の一部を表1に示す。これらを通じてOKIグループ全体に開発実績を蓄積すると共に、これらの実践にあたっては教育啓蒙と情報共有コミュニティと密接に連携している。まだ道半ばではあるが、こうした事例をプレスリリースなどで積極的に発信し、今後も継続して認知度向上を図っていく予定である。

表1 活動事例

事例	概要
本人確認書類の偽造判別支援	本人確認書類(運転免許証)の偽造可能性の判定
手書き帳票認識	手書き帳票の自動デジタル化支援
感情推定	人の潜在的な感情の推定 ¹⁾
信号ノイズ除去	センサー信号の高度ノイズ除去 ²⁾
音響認識	音響信号を用いた状況推定
印刷品質評価	印刷品質の定量的な自動評価

まとめと今後の展望

本稿では、「ディープラーニング開発エコシステム」と呼ぶ活動の概要と、その三つの活動の柱を説明した。

OKIは、通信・運輸・交通インフラ・金融領域などで多様

な商品群があり、それらは多様なデジタルデータの流通や分析を担ってきた。これらの領域で得られた数多くの事業知識は、多様なAI技術を適用することにより、多くの価値に変化し、より良い社会の実現へと活用される。OKIは、こうした豊富な事業知識を活かし、ディープラーニング開発エコシステムにより高められた技術力を合わせることで、単に技術的な定石を投じた課題解決ではなく、それぞれの事業領域に応じた有効で安全なAIを活用するAI-Readyな社会の実現に貢献していきたいと考えている。このため、ディープラーニング開発エコシステムの活動も現在のパラダイム変化への適合を目指したのから、AI-Readyな社会実現に向けたOKIのAI価値を創造し、共有し、そして、その質を高めていく活動へと変化させていきたい。

■参考文献

- 1) OKIプレスリリース ディープラーニングを用いた「感情推定技術」の研究開発を開始、2018年4月3日、
<https://www.oki.com/jp/press/2018/04/z18001.html>
- 2) 根本亮、清賀史暁、越川博昭、橘素子、山本康平:シミュレーションを活用したディープラーニング技術開発とエッジ端末への実装、OKIテクニカルレビュー第233号、Vol.86 No.1 pp.28-31、2019年5月

●筆者紹介

橘素子:Motoko Tachibana. 経営基盤本部 研究開発センター イノベーション推進室
前野蔵人:Kurato Maeno. 経営基盤本部 研究開発センター イノベーション推進室

TiPo 【基本用語解説】

PoC(Proof of Concept)

新しい概念や理論、アイデアの実現可能性を示すために、簡易な試作品などを作って検証すること。

API(Application Programming Interface)

アプリケーションを開発する際に、共通的に利用できる機能呼び出すことが可能な仕組み。便利な機能をあらかじめAPIとして提供することで、開発者が一から作らずに短期間でアプリケーションを構築できる。