

# AI を含むソフトウェア開発の進め方についての研究

松山 和弘

2020 年 10 月 8 日

## 1 研究計画

今後、AI を含むソフトウェアの開発が増えていくと、AI についての過度の期待や要求、開発関係者の理解不足のために、実現不可能な機能や品質を要求仕様として設定し、誤って開発を進めてしまうことが起きると考えられる。

開発時や運用時に遭遇する問題をいくつか挙げると、以下がある。

1. 「計算結果があらかじめわからない機械学習」を既存の手法で開発してしまう問題  
「計算結果があらかじめわからない機械学習」を、既存の「計算結果があらかじめわかるコードの動作」の開発手法で開発を進めてしまうことが起こり得る。さらに、実際のソフトウェア開発は、「計画駆動型プロセス (ウォーターフォールモデル)」で行われることが多く、AI を含むソフトウェアについて実現不可能な機能や品質を要求仕様として設定したまま、変更されずに開発が進んでしまうことが起こり得る。
2. 機械学習の劣化問題  
ソフトウェアの開発段階の機械学習では品質要件を満たしたとしても、実運用段階では品質要件を満たさないような状況が発生すると考えられる。  
例えばコロナウイルス前と後では、人間の行動が変化しているため、コロナウイルス前の学習内容は、コロナウイルス後には適合しないかもしれないが、要件未達として扱われるかもしれない。
3. 学習データのブラックボックス化問題  
データ学習型 AI のうち、SVN、デブラーニング等は、学習用に使用してした元

データと切り離れた学習結果のデータを人間が見て、動作を推測することは、実際に動作してみない限り困難である。

例えば、開発に入手した学習済データ（例えば画像認識のための深層学習済データ）や、前処理済の深層学習に使用するデータ、他のベンダが提供する画像認識などのサービスは、プライバシー侵害や人種差別などの人権侵害に関するデータを元に学習したかもしれないが、ブラックボックス化したデータは、そのデータだけ見ても、適性なデータかどうかわからない。学習済データの利用者や開発者は、無自覚のうちに不正なデータを使用して、不正行為に加担しているかもしれない。

これらの、AI を含むソフトウェア開発時に遭遇する問題に着目し、研究方法は以下とする。

1. 類似の研究、開発事例を調査する。
2. AI を含むソフトウェアが、既存の開発プロセスに混入してしまう場合の問題を考察する。
3. AI を含むソフトウェア開発を、誤って進めない仕組みを検討する。例えば、「計算結果があらかじめわからない機械学習」と、既存の「計算結果があらかじめわかるコードの動作」部分の仕様記述の分離方法について検討する。

## 2 志望理由

現在私は、メーカーにて、ソフトウェア開発業務に携わっている。これまで、製品向けの OS、コンパイラ、ミドルウェア等の開発や保守を担当してきた。

これまで担当してきた開発は、主に「計画駆動型プロセス」で行うことが多かった。

「計画駆動型プロセス」では、開発開始前に要件と開発規模を定めるが、実際の開発現場では、要件が曖昧のまま開発を始めることは一般に多く行われていると考えられる。開発中の要件変更により、開発規模の増大したとしても、開発を完了できれば、開発は成功扱いになっていたと推測できる。

今後、「AI」が開発要件に入ってくると考えられるが、開発現場で一般に行われている「計画駆動プロセス」等のこれまでの開発や品質管理手法をそのまま当て当てはめて開発を進めた場合に、開発現場の「カイゼン」だけでは立ち行かなくなると危惧される。このため、「ソフトウェア工学」の観点で「AI」の扱いを見直す必要があると強く感じている。

以上を踏まえて、貴大学院情報学プログラムを志望した次第である。