

クラウド・ビッグデータとソフトウェアエンジニアリング

鯨坂恒夫
和歌山大学

ajisaka@sys.wakayama-u.ac.jp

要旨

クラウド・ビッグデータと高頻度で抱き合わせにされるこれら二つのバズワードについて、その本質をさぐりながら、ソフトウェアエンジニアのとるべき姿勢と態度について議論する。

1. はじめに

クラウドとビッグデータは、今後も当面は情報システム技術のトレンドであろうということが共通点であり、現象としても共起する傾向にはあるだろうが、対応すべき技術的内容は必ずしも相似していない。したがって、議論展開は実はオムニバスのようになる。

2. クラウドによる変化

利用者と開発者のギャップはこれまでも様々に言及されてきたが、クラウドは両者のものの見方をさらに分離する。利用者は使いたい機能を使いたいときだけというスタンスになるのに対して、開発者は情報統合の視点を強化しなければならなくなる。利用者側について運用技術を支援する役割と、クラウドシステム自体を開発する役割に明確に分かれる。これまでソフトウェアエンジニアが一人二役で両方を担ってきたが、頭の使い方やネゴシエーションのしかたが相当に違うので、担当者が分かれる方向に向かうと予想され、それによって利用者・開発者ギャップ問題が思わず解決するのかもしれない。

利用者側の技術も開発者側の技術もこれまで以上に業務の内容とプロセスに密着することになる。利用側は業務とシステムとのマッチング(機能選択と連結)である。エンジニアは対象業務のビジネスモデルやワークフローをよく理解していなければならない。エンドユーザコンピューティングの指南役も求められるかもしれない。一方、開発側のほうは、情報系クラウドが主流の現状ではその業務結合性は明確でないかもしれないが、基幹系クラウ

ドがこれから進展してくれば、情報の要素認識と意味的關係性の把握が重要になってくる。単体業務ではなくグッズとサービスのチェーンマネジメントを見通していなければならず、それによる規模と複雑さ・不確定性の拡大に対抗する手だてが肝要となる。

空想的にはクラウドがどんどん結着して「ザ・情報システム」がひとつあればよい。組織活動にはすべからく inbound logistics と outbound logistics がある (cf. Michael Porter の Value Chain, 1985) ので、その鏡像關係がクラウド内で縮退すれば、全体としての規模と複雑さは半減する。問題は不確定性である。規格化に人々が反抗するというのではない。社会の進歩により概念要素や規則が変化するというのである。ソフトウェアの問題解決はパターン化しかないが、進歩によって生じるバリエーションを抑圧することはできない。この二律背反を解消するのはバリエーションの生え方をパターン化する技である。パターンは還元論的であり、バリエーションは全体論的であるので、これら不倶戴天の両者を協調させる超絶技巧が求められる。

ザ・情報システムが空想であるとするれば、空想から科学へ移行するには何が必要かを考えることができる。ひとつ明らかなことに、ザ・情報システムは一日にしては成らないので、そこに至るステップはいかにあるべきかという課題がある。あるいは IoT がそこへ向かうのを加速するかもしれない。何か多くの局面で有用な thing がネット化されれば、それを共有しようといういろいろなサブシステムが寄り集まり融合をはじめ。

3. ビッグデータというブルートフォース

クラウドはソフトウェアエンジニアリングや情報システム技術のコンテキストから出てきた由緒正しいトレンドだが、ビッグデータは必ずしもそうではない。クラウド化すれば扱うデータ量も巨大化するので、ビジネスデータという限られたジャンルのビッグデータが連想されるというだけである。ビッグデータはむしろ WWW やセンサネットワーク、IoT から溢れ出てくるものが主流であって、その処理のさ

れかたはかなり武骨である。ビッグさは半端でないので、精緻な扱いはもとよりできない。何種類かの特徴値の組合せを分類するだけである。言語処理になぞらえていえば、構文解析以上の高級な処理は一切なし、字句解析のレベルのみという、すでにインターネット検索で成功した手法である。

そうして得られた結果(予測)には理由がない。Singularity (cf. Kurzweil, 1999,2005)をもたらすという「新型人工知能」はこの方式である。これはコンピュータの本分をわきまえたという点では妥当であり、それによってかつての(人間を模倣しようとした)人工知能の失敗を乗り越えている、というより別物である。そこで人間が why の追求をやめるか続けるか、いいかえれば「知性」の定義そのものを変えるか変えないかによって、人間の尊厳の維持・存続が決まる。

ビッグデータの処理技術は特徴値の種類を選択と各特徴値の閾値の設定にかかっている。さらにソフトウェアがそれらを動的に最適化できることが必要だと思われるが、こういうメタサーキュラーなしくみはなかなかうまく動かない。したがって、得られた予測を見た人間の対応行動が早ければ(例えば金融商品売買や混雑状況対応など)、次の予測は狂うか大きく変動して、意思決定の拠りどころにならなくなる。新型人工知能に人間の理解する意味に基づいた調整を混ぜると、おそらくそれ本来の威力を発揮しなくなるので、未来予測が際限なく進歩することはないと思いたい。

4. おわりに

クラウド・ビッグデータとソフトウェアエンジニアリングという標題を受けた結論は何かといえば、不確定性(uncertainty)への対応強化である。これまでソフトウェア工学論文の前振り定番であった大規模化・複雑化は、分析・設計手法の進歩とプラットフォーム・フレームワーク・パッケージ化技術の進展により一定の対応がなされてきた。ソフトウェアの生産性と品質を向上させる決定打は、できるだけソフトウェアを作らない(再利用する)ことであるという信条が浸透してきた。しかし、これまでも保守の問題の多くが単純な修正保守より適応保守・改善保守の要求であろうと思われ、クラウド化によりますますその傾向は強まる。このように不確定性の増えた状況でも、それに対応していちいちソフトウェアを作らないですますには、先述のとおりバリエーションのパターン化という形容矛盾を実現しなければならない。

これはおよそ人間の活動とはどのようなものかという、

神の発するような質問にこたえようとするものである。行動観察のようなボトムアップ的手法や、Luhmann の社会システム理論、Latour の Actor-Network Theory, Vygotsky に始まる Activity Theory, Bertalanffy の一般システム理論など、社会学的・哲学的論考もあるが、どれもまだ大成功しているようにはみえない。しかし、相手にしているのが全体論的枠組みであるから、ゲーデルの不完全性定理のような明解な否定的解決も確立するとは思えず、挑戦を続けるしかない。

上記「社会学的・哲学的論考」に関する補足:筆者も原書をあたってきたわけではありません、ネットワーク経由でかじったものです。ニクラス・ルーマンについては最も日本語の情報や訳書が豊富ですので、容易にいろいろ参照できます。Actor-Network Theory と Activity Theory については英語版 wikipedia がしっかり書けていると思われます。ベルタランフィについては、「松岡正剛の千夜千冊」(<http://1000ya.isis.ne.jp/0521.html>)からあたり始めるのも一興かもしれません。