



NTTコンピュータ&データサイエンス研究所
特別研究員

齋藤 忍 Shinobu Saito

ソフトウェアのゴミを削減し、持続可能なIT社会を実現する「ソフトウェアデジタルツイン」

2022年の世界のITサービスのマーケット規模は1.3兆USドル(約150兆円)と報告されており、市場の要求によって生み出された膨大なソフトウェアは、私たちの生活に大きな価値をもたらす一方で、さまざまな問題を引き起こすことも指摘されています。これらの問題の解決に向けて、ソフトウェアを増やす(生産)ことに目を向けるだけではなく、ソフトウェアを減らす(減量)ためのアプローチが期待されています。今回は「ソフトウェアデジタルツイン」の研究に取り組む、齋藤忍特別研究員にお話を聞きました。

◆PROFILE: 2001年慶應義塾大学大学院修士課程修了。同年、NTTデータ入社。2007年慶應義塾大学大学院博士課程修了(工学)。デジタル時代における新たなソフトウェアエンジニアリングの研究に従事。2022年度情報処理学会コンピュータサイエンス領域功績賞、APSEC2022 Best SEIP Paper Awardなどを受賞。2016年~2018年カリフォルニア大学アーバイン校ソフトウェアリサーチ研究所 客員研究員。2018~2021年NTTソフトウェアイノベーションセンター 特別研究員。2021年よりNTT コンピュータ&データサイエンス研究所に所属。



ソフトウェアの3R (Reduce, Reuse & Recycle) で、ユーザ・開発者にとって理想的な環境を実現

■「ソフトウェアデジタルツイン」とはどのような技術でしょうか。

私が研究している「ソフトウェアデジタルツイン」とは、ソフトウェア内部の状況をコンピュータ上のデジタル世界で再現することで、ソフトウェアの見直しや削除すべき箇所(プログラム)を発見・予測する技術です。近年のソフトウェアを取り巻く状況として「ソフトウェアのゴミ」と呼ばれる、全く活用されていないソフトウェアの増加が大きな課題として指摘されています。本来、より良い世界を実現することを目的として生産されたソフトウェアが、結果的にさまざまな問題を引き起こしているのです。これらの問題はビッグコード(Big code)やソフトウェア膨張(Software bloat)と呼ばれています。前者は、ソフトウェアが大きくなり過ぎた結果、限られた人(開発者)では維持が難しくなっていることを意味しており、いずれソフトウェアの量はプログラマの能力でカバーできる範囲を超えると予測されています。実際に海外の調査では、半数以上のプログラマが「10年前と比較して扱うプログラムのコード量が100倍以上に増えている」と回答しているほどです。そして後者は、増えすぎた(膨張した)ソフトウェアがCPUやメモリといったシステムのリソースを大量に消費してしまうことを意味します。その結果として、たくさんの電力消費とCO₂排出をもたらす、地球環境にも悪影響を与えてしまいます。

私自身はNTTデータに入社してから、「ユーザにとって良いソ

フトウェアを開発者がラクにつくるためにどうすればよいのか」を長年考えてきました。しかし数年前、ある経済誌に書かれていた「企業で活用されずにユーザに無視されているソフトウェアはゴミである」「平均的な企業のIT予算の3分の1はゴミに浪費されている」という記事を見て大きな衝撃を受けました。確かにこれまでも「開発されたシステムの機能の半分以上はほとんど使われていない」といった報告もありましたが、「私の研究が結果として、人や地球の環境に悪影響を与えている可能性もあるのではないか?」という想いを抱くようにもなり、現状を打破するアプローチを考えるようになりました。

そのようなことを考えていたある日、小学生の子どもの夏休みの宿題を手伝っているとき、子どもの教科書に3Rという環境問題にまつわるコンセプトが紹介されていました。3Rとは「Reduce(発生抑制), Reuse(再使用), Recycle(再生利用)」の総称です。これを見て私は「ソフトウェアを生産してばかりのITの社会にも、このコンセプトを当てはめていくべきでは?」と思うようになりました。今後も人々の暮らしを豊かにするソフトウェアを「つくる」ことは必要である一方、作り過ぎた、増えすぎたソフトウェアを「なくす」ことも人や環境にとっては重要になると考えたからです。そのような着想から生まれたコンセプトが「ソフトウェア3R」です(図1)。この「ソフトウェア3R」ではソフトウェアをつくるだけではなく、不要なソフトウェアを減らすことや既存のソフトウェアを使い回すことを推進しています。これによって、ソフトウェアのユーザや開発者にとって持続可能なIT社会を実現できると考えています。昨年、ソフトウェア工学の国際会議でこ

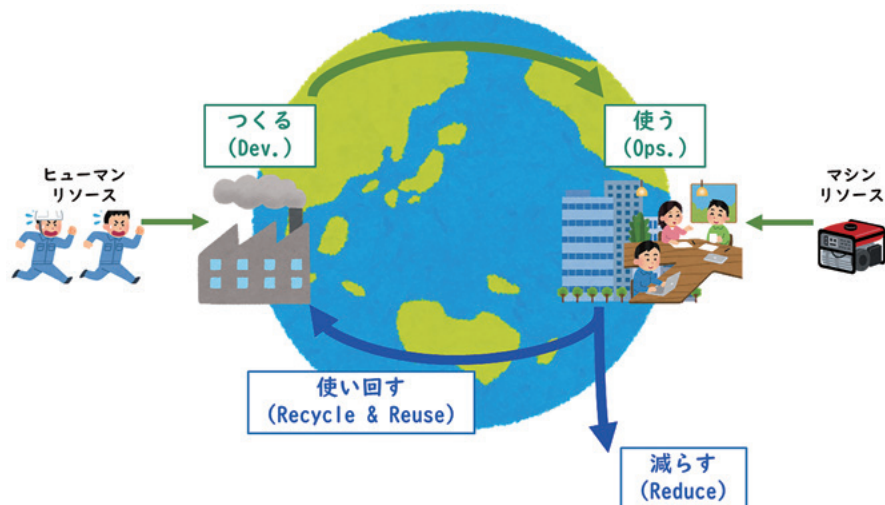


図1 ソフトウェア3Rが実現する世界

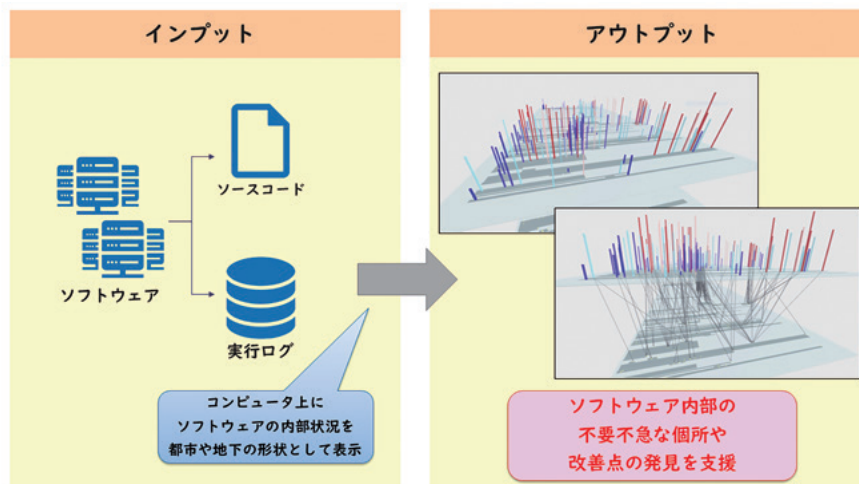


図2 ソフトウェアデジタルツインの概要

のコンセプトを発表したところ、参加されていた研究者の方からも面白いと言ってもらえました。

■具体的にどのようなアプローチで研究をされているのでしょうか。

「ソフトウェア3R」のコンセプトを実現するアプローチとして、現在研究を進めているのが「ソフトウェアデジタルツイン」です。これは都市や人体など現実世界の対象から収集したデータをコンピュータ上のデジタル世界で再現する「デジタルツイン」の研究から着想を得ています。例えば人体のデジタルツインでは、現実

に動いている脳や心臓を直接観察することが困難なため、それらをデジタル空間上に表現して、病気の原因究明や将来の病気の予測とそれに対する予防等に活用する研究などが進められています。ソフトウェアもたくさんのプログラムの集合体として構成されており、動いているソフトウェアの内部の状況を直接観察することが難しい対象であるといえます。そのため「よく使われているプログラムはどこか」「逆に使われていないのはどこか」といったことを正確に把握することは困難を極めます。そこで「ソフトウェアデジタルツイン」では、ブラックボックスのソフトウェアの内部状況を高精度に再現し中身を「丸見え」にすることで、不要不急や改善すべき箇所（プログラム）の発見・予測をめざしています（図2）。

この研究はNTTが提唱するIOWN (Innovative Optical and Wireless Network) 構想において、デジタルツインコンピューティング (DTC : Digital Twin Computing) とのかかわりが非常に強いです。現在のDTCでは都市や医療といった現実 (リアル) の世界を対象とした領域で検討が進んでいます。この「リアルの世界をサイバーの世界に再現」するアプローチに対して、私の研究では「サイバーの世界を別のサイバーの世界に再現」、つまり「コンピュータ上で動いているソフトウェアの内部を別のコンピュータ上に再現」しようとしています。サイバーのものを別のサイバーのものとして再現するというのは一見すると不思議かもしれませんが、これによって初めてソフトウェアの内部を人が正しく理解することができ、予測やシミュレーションが可能になると考えます。

ソフトウェアの実行状況を高精度に再現するために扱う主なデータは、プログラムが記述された「ソースコード」と、プログラムの動作内容を記録した「実行ログ」の2種類です。後者の実行ログの収集が、研究を進めるうえでも実務で導入するうえでも大きなハードルです。理想は、ソフトウェア内部の「すべての個所 (プログラム)」の動作の記録を、ソフトウェアが実行されている「すべての期間」において収集することです。しかしそのためには、すべての個所の動作を記録するための新たなプログラムを人が作成する必要があります。またすべての期間において実行ログを収集しようとする、システムのリソースも余計に消費してしまいます。これでは本末転倒で、人にも環境にも悪影響を与える結果となってしまいます。

そこで現在取り組んでいる研究では、ソフトウェア内部の限られた個所において限られた期間で得られた実行ログのデータから、ソフトウェア内部の全体の実行状況を推定したり、予測したりするための手法を検討しています。大学との共同研究などを通じて、統計分析や機械学習を応用することで手法の確立をめざしています。

研究を基礎研究で終わらせないために、「孤独」と「交流」の営み続ける

■今後の研究ビジョンについて教えてください。

ソフトウェアデジタルツインにおけるこれまでの研究成果の1つとして、業務システムの各機能の利用頻度や、システム全体のオペレーションを3次元で可視化する技術「ProcessCity (プロセスシティ)」があります。すでにNTTの複数の事業会社に成果を提供しており、昨年のNTT R & Dフォーラムでは提供先の1つの会社と共同出展や報道発表もしています。また別の提供先の会社と共著で書いた論文は、昨年開催されたソフトウェア工学のアジア・太平洋地域で最大規模の国際会議においてベストペーパー

として表彰されました。

現在、レガシーシステムと呼ばれる老朽化・複雑化した古いシステムを維持するための費用が、企業のIT予算の6~7割にのぼるともいわれています。またそれらのシステムから縮小・廃止する機能を発見するための調査は、期間が何カ月にもわたるケースがあります。そこで本研究の成果を活用することで、まずは調査期間の短縮化や調査内容の精度向上を実現して、調査段階の工数を2~3割削減、最終的には10分の1にまで減らすことをめざします。

2022年の世界のITサービスのマーケット規模は1.3兆USドル (約150兆円) というレポートがあり、先ほどの経済誌の記事によれば、これらのITサービスの約3分の1である50兆円がゴミとして浪費されている試算になります。もちろんNTT単体で世界中のソフトウェアのゴミを減らすことは難しいでしょうが、国家予算の規模にもなるこれらの莫大なゴミを削減するための技術を創出し、NTTの顧客企業などへの大きな価値提供につなげていきたいと考えています。

さらに今後のITサービス業界では、「ソフトウェアを安心・安全に解体するための産業が形成されてくるのではないかと」予想しています。これは、ソフトウェアを生産するのが建設業だとすれば、ソフトウェアを取り除く (削除する) のは解体業であるという考え方です。今後も研究を進めることで新たな産業の形成や事業機会の創出につなげ、NTTのコア事業領域の1つとしてソフトウェア解体業が新たに追加されることが私の究極の目標です (図3)。

■NTT 研究所にはどのような印象をお持ちでしょうか。

私が所属しているNTTコンピュータ&データサイエンス研究所では、名前のとおりコンピュータとデータに関する研究をしており、あらゆる業界業種にわたる多種多様な研究に取り組んでいます。その中でグループ企業内にソフトウェアをつくる企業も使う企業も存在していて、ユーザと開発の両方のフィールドがあることは非常に強みだと感じます。ステークホルダの間口が広く、私も実際に研究を進めるうえでソフトウェアを使う立場のグループ企業と意見交換できる場を設けていただいたこともあり、とても恵まれている環境だと思います。

またグループ外の企業の方々とお付き合いする際には、NTTのネームバリューを実感する場面も多いです。例えば10年前にある国際会議に初めて参加したとき、誰も知り合いがいなかったため、受付でネームプレートをもって1人で会場まで歩いていると、隣を歩いていた方が私のネームプレートを見て、「NTTなんだ、〇〇さんを知っているよ」ととても気さくに声をかけていただいたことがあります。その方はソフトウェア工学の世界ではとても

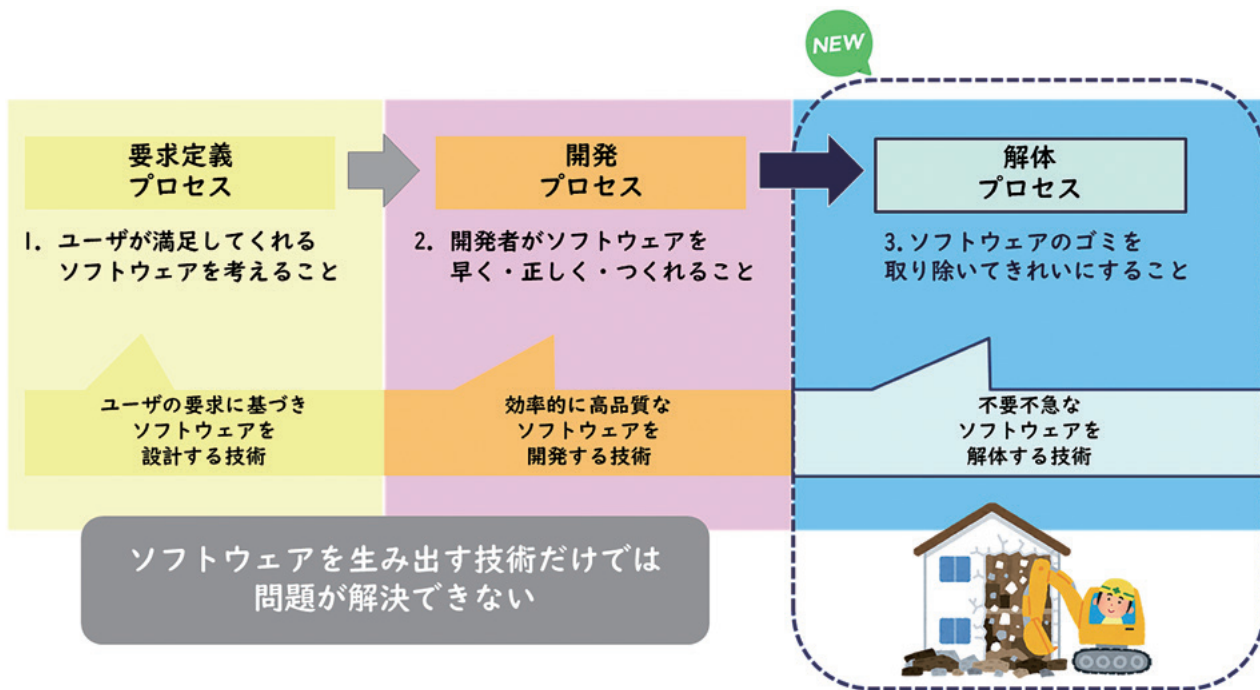


図3 ソフトウェア開発における解体プロセス

有名な先生で、その出来事がきっかけで今でも国際会議でお会いするとお話しして、研究についての相談をするなどのつながりがあります。もちろんNTTのネームバリューがあるからといって必ずしも研究がうまくいくということではありませんが、自分の知らない環境や場所に行く際には非常に心強いものであると感じています。

■最後に研究者・学生・ビジネスパートナーの方々へメッセージをお願いします。

NTTで研究を進める中で、私は「基礎研究で終わる研究はない」という言葉を信条にしています。これは私がNTTデータに入社したころ、研究所から自部署に来られた方に教えていただいた言葉です。その意味は、「すぐに用途が見込めないような基礎的な研究であっても、その研究が何の役に立つのかという意味が必要である。意味がない（語れない）のであればそれは基礎研究ではなく、そもそも研究でもない」ということでした。私はこの言葉をいつも心にとめて、研究費をかける以上は研究成果が何の役に立つのかといった「研究の意味」を提示することを心掛けています。

同じ時期にやはり先輩社員の方から聞いた「本来、研究は“孤独”なものである」という言葉も記憶に残っています。研究はオリジナリティが求められるため、孤独というのはある面では正しいです。私自身の研究を振り返ってみても、基本的なコンセプト

を考えていく段階は、自分でうまく言語化ができず、1人で悩み続けることが多いと思います。しかしそのコンセプトを実現する段階では、多くの方の協力を得ることが必要です。NTTでは他者・他社と「交流」ができる環境や機会もあるので、分野外の新たな知識を得たり、自分1人では到達できない新たな世界が見えたりすると、とても高揚（ドキドキ）することも多々あります。今後も強い気持ちと信念を持って研究に取り組み、自分に足りないものを多くの皆さんと交流をしながら一緒に研究ができればと思っています。



(今回はリモートにてインタビューを実施しました)