

若手エンジニアの育成と輩出を目的とする サイボウズ・ラボユースが創立10周年

教育用 OS の xv6 に関する研究など、
これまでで最多となる第10期生15名の研究成果を発表

サイボウズ株式会社（本社：東京都中央区、代表取締役社長：青野慶久、以下サイボウズ）は、サイボウズグループの研究開発部門であるサイボウズ・ラボが運営する若手エンジニア育成プログラム「サイボウズ・ラボユース」が創立10周年を迎えたことを発表いたします。



■設立の背景

「サイボウズ・ラボユース」は、世界に通用する日本の若手エンジニアの発掘と育成を目指すことを目的とし、学生の若手クリエイターに研究開発の機会を提供する場として、2011年3月31日に設立されました。同プログラムでは、日本全国より応募のあった高校生、大学生、大学院生、バグハンターの中から対象者を選抜し、サイボウズ・ラボの社員が直接、開発を指導し、それぞれの研究テーマへの取り組みを支援します。研究成果の権利は学生に所属し、サイボウズ・ラボユースは純粋に、サイボウズの技術者サポートのもと、社会に優秀な若手エンジニアを輩出することをその目的としています。

■サイボウズ・ラボユースの特徴

サイボウズでは、社員一人ひとりの「やりたいこと」「できること」「やるべきこと」の3つを「モチベーション3点セット」と呼び、それぞれの成長を支援しています。サイボウズ・ラボユースでもその理念のもと、学生の自主的なモチベーションを大切にしています。そのため、研究内容についても学生自身が自由にそのテーマを選択することが可能です。また、サイボウズ・ラボユースは、通常の奨学金制度とは異なり、研究成果の権利を学生自身のものとし、卒業後の進路等についての規定も定めていません。卒業生



は、当社に縛られることなく、自由に卒業後の進路を選択することが可能です。

学生の研究開発に際しては、サイボウズ・ラボの各エンジニアがそれぞれメンターとしてバックアップし、全面的にその研究を支援します。

■これまでの研究生とその進路

サイボウズ・ラボユースはこれまでに、秋田大学工学部で助教を務め、著作として『正規表現技術入門』（技術評論社）を発表する新屋良磨（しんやりょうま）氏、東京工業大学学術国際情報センターで助教を務める石井将大（いしいまさひろ）氏、経済産業省の未踏スーパークリエイターに認定された城倉弘樹（しろくらひろき）氏・橋本論（はしもとろん）氏・怒田晟也（ぬたせいや）氏をはじめ、WebKit の開発に関わる鈴木勇介（すずきゆうすけ）氏、Google で活躍する hikaliium 氏、さらに現在サイボウズ・ラボ社員でもあり『ゼロからの OS 自作入門』（マイナビ出版）を執筆した内田公太（うちだこうた）など、官民に渡り多様な人材を輩出しています。

10 周年に際し、卒業生の方々からも以下のコメントをいただいています。

新屋良磨氏（秋田大学工学部 助教）

「サイボウズ・ラボユース制度誕生から 10 周年と聞いて感慨深いものがあります。私は修士一年の頃、ラボユース第 1 期生として「正規表現 JIT エンジンの実装」というプロジェクトに取り組んでいました。東工大のある大岡山と当時サイボウズ・ラボのオフィスがあった飯田橋を南北線で往復していたあの頃から 10 年経ったことになりましたが、今でもラボユースの記憶は色濃く残っており、その後の自分の進路にも大きな影響を与えてくれました。

当時はプログラムの高速化にハマっていて、プロジェクトを始めた段階では修士号を取ったあとはプログラマーになるつもりでいました。チューターの光成さんは高速化の達人で、ラボユースを過ごした 1 年間にプログラム高速化のいろはを叩き込んでいただき、そのおかげもあって最終的には高速化の成果・技法を学術論文にまとめることができました。

プログラム高速化について引っ張られるように理解と実装力が高まった 1 年でしたが、その一方で高速化に対して自分の中で「やりきった感」を感じました(もちろんまだまだ高速化の余地はあるのですが)。

また、ラボユース同期生の鈴木さんとはラボから帰る方向がたまたま同じで、南北線の中でプログラミングの議論を良くしていましたが、彼のプログラミングに対する知識や情熱はものすごく、同年代で初めて「(プログラミングでは)こいつには勝てないな」と思わせてくれた人でした。

以上のような背景もあり、ラボユースでの 1 年間を終えた頃から、私の興味は高速化から正規表現の理



論の方向に徐々に傾倒していきました。

チューターの光成さんが数学科出身ということで、数学の雑談や正規表現の数理的背景の話にも付き合ってくれた点も、興味の転換に大きく影響していると思います。実際、ラボユース以降は正規表現の理論研究に取り組み、プログラマーになることを辞めて博士課程に進み、現在は正規表現の研究者として活動しています。

奨学金をいただいて自分で研究を主体的に進められただけでなく、チューターや他のユース生からいろいろ「熱い」影響を受けたことが私にとっては最も重要なラボユースでの経験でした。

私や過去のユース生のように、今後もラボユースで「熱い」影響を受ける方々が続々と現れることを期待しております。ラボユース 10 周年、あらためておめでとうございます。」

城倉弘樹氏 (LINE ITSC Verda Network Development Team)

「ラボユース 10 周年。おめでとうございます。

当時の一年間の経験は現在の私のソフトウェアエンジニアとしての重要要素である『設計/開発能力』に大きな影響を与えてくれました。

特に当時は数千~万行規模のソフトウェアをフルスクラッチで開発したことがなく、C/C++での実践的な設計テクニックやマルチスレッドプログラミングの TIP などをメンターの光成さんから学ぶことができたのはとても貴重な経験でした。

現在私は LINE のサービスを収容している Cloud のコアメカニズムの企画/設計/開発を行っています。そのようなフィールドでも『ひとまずこのイメージで作ってみよう』『このような設計方針で開発する』といった、反射神経は非常に重要です。

ラボユースでの経験はこのようなスキルを研ぎ澄ますための修行であり、重要な通過点であったことを確信しています。

また、『新規性を出すことに執着せずに作りたいものを存分に作る』という方針を尊重してもらえたことも重要でした。

近年は多くのソフトウェアが非常に早いスピードで進化していき、『New』を作り出すのは簡単なことではありません。

そして、『筋の良い車輪の再発明』をすることによって多くの知識を習得し、将来的な『New』を生み出すきっかけにできるということはいうまでもないでしょう。

そう言ったトピックを考えた時に、ラボユースでは『まず手をうごかして開発を試みる』という取り組みができるのは良いことだと思います。

今後のラボユースのメンバー、サイボウズ・ラボのご発展を心より期待し、応援しています。」

田辺敬之氏（ノーチラス・テクノロジーズ）

「初めまして。サイボウズ・ラボユース 8 期生の田辺敬之と申します。

私はラボユースの活動として、トップ国際会議にて提案された並行性制御法を中心に複数手法を再設計・再実装・評価を行いました。

それを用いて実験・分析論文を執筆し、データベース系において世界最高難度の論文誌である PVLDB に論文を通しました。

師事したメンターがテーマに関連する識者であったこと、私が苦学生であったことから、ラボユースが無かったらトップ会議に論文を通すことは叶わなかったと思います。

ラボユースは自由にテーマを選択でき、業界に名を馳せるメンターらへ気軽に助言を求めることができ、議論ができて、奨励金を頂ける素敵な環境です。

情熱を注いで活動したい対象がある人は、ぜひ応募を検討してみてください。」

■第 10 期生の研究成果

3 月 30 日に「第 10 期サイボウズ・ラボユース成果報告発表会」がオンラインで開催され、下記の通り、第 10 期生の研究成果が発表されました。教育系 OS である xv6 に関する研究など、節目となる設立 10 周年を迎え、サイボウズ・ラボユースとしても新しい取り組みについて発表されました。



「第 10 期サイボウズ・ラボユース成果報告発表会」の様子

それぞれのゼミおよび担当メンターと、卒業生の研究発表に関するコメントは以下の通りです。

■ 低レイヤゼミ（担当：内田 公太）

鈴木 友也 「xv6 を自作 VMM 上で動かす」

「xv6 という教育用 OS を自作 VMM で動かすというテーマで話します。私のメンターでもある uchan さんがみかん本を出版されたこともあり、最近は自作 OS 界隈の盛り上がりをととても感じます。

自作 VMM は必要とされる知識が自作 OS と近く、それでいて意外と挑戦者はまだまだ少ない領域です。今回の発表で自作 VMM に興味を持ってもらえたら嬉しいです。

発表では、ラボユースの活動期間で実装した kvm という VMM を取り上げながら、その構成や苦労した点・よかった点などを振り返りつつ解説します。

kvm はゲスト命令の実行レイヤとして KVM を利用し、アプリケーションレイヤでデバイスや割り込みのエミュレーションを行うホスト型の VMM です。類似のソフトウェアとして、QEMU などがあります。」

徳永 大貴 「USB デバイスドライバの実装」

「ラボユース・低レイヤゼミにおいて、USB キーボード、USB マウス、USB フラッシュメモリのドライバをフルクラッチで作成しました。作成において基盤となる OS は、すでに自作中であった Ramen OS を使用しました。使用言語は Rust です。

具体的な開発内容として、PCI デバイスからホストコントローラの探索、ホストコントローラの初期化、USB 機器の接続状態取得、接続されている機器に対する初期化、デバイスの種類に応じた処理の実行のコードを記述しました。

今回の発表では、USB フラッシュメモリのドライバに焦点を当てました。USB フラッシュメモリに対しては、データの読み出しと書き込みの処理を実装しました。このために SCSI プロトコルの実装を行いました。」

花井 一輝 「xv6 にリモートファイルシステム 9P を実装する」

「xv6 にネットワークスタック、9P プロトコルを実装し、Linux/xv6 間でのファイルシステムの共有を実現したことについて話します。

xv6 とは、MIT で開発されている教育用のコンパクトな OS です。メモリや CPU、ディスク等のハードウェアリソースに加え、プロセスやファイル等の基本的なリソース管理をサポートしていますが、公式にはネットワークの実装がありません。

私達は普段からインターネットを利用しており、モダンな OS にとってネットワークの機能は必須なものです。そこで、xv6 に対して TCP/IP のネットワークスタックを実装しました。更に、そのネットワークスタック上に、9P と呼ばれるファイルシステムプロトコルを実装し、Linux/xv6 間でのファイルシステムの共有を実現しました。

今回の発表では、技術的な解説を含めた labyouth での取り組み、実装するにあたり苦労したことや力を入れた点等を説明します。また、発表の最後にはデモも行います。」

■ 言語処理系ゼミ（担当：川合 秀実）

菅原 大和「複数アーキテクチャを対象とする自作言語コンパイラ的设计と x86_64 ツールチェーンの開発」

「コンピュータサイエンスの学習目的でシンプルなコンパイラを実装するという例は多く存在しますが、その殆どが汎用 PC に搭載されている CPU アーキテクチャの x86_64 を対象としたものです。

これには、"フロントエンドとバックエンドに分かれている既存コンパイラとのギャップ"や、"x86_64 に代表される CISC アーキテクチャの知識に偏ってしまう"という問題があると考えます。

これらの状況を鑑みて、本活動ではスタックマシンを用いたシンプルなバックエンド、かつ複数アーキテクチャに対応しうる設計をモットーに自作言語のコンパイラを実装しました。

また、コンパイラの実装に合わせて、x86_64 のツールチェーン開発を行いました。

最後に、elf-utilities という ELF バイナリを扱いやすくするライブラリを開発しました。

実際に elf-utilities を用いて elfpeach という TUI 版の readelf ツールのプロトタイプを作成しています。」

飯田 圭祐「静的型付きスクリプト言語の開発」

「言語処理系ゼミで、2年半前から開発を続けていた静的型付けの自作スクリプト言語 maxc のフルスクラッチ実装を行いました。

maxc では、人間が書きやすく、あまり見た目が煩雑にならない文法を目指しました。

発表では、maxc の基本的な機能や実行を高速化するために行った最適化について発表しました。」

(<https://github.com/k-mrm/maxc>)

松井 誠泰「OLAP SQL クエリの静的解析と自動並列化」

「機械学習やデータ解析などの技術は強力で柔軟なシステムを実現するがデータと処理の依存関係やデータ・処理それぞれの変更への追従など管理の難しさがあります。

そこで型・スキーマの安全性を保ちながらデータ基盤の構築を行うための AlphaSQL というフレームワークを提案します。AlphaSQL は、SQL ファイル間の依存関係を解決することにより、SQL ファイル集合全体の型・スキーマ解析と並列化を行います。テーブルを作成する SQL 文に対する、作成されたテーブルを参照するクエリの依存関係が、SQL ファイルの静的解析によって自動的に抽出されます。結果は有向非巡回グラフとして出力され、ユーザーは依存関係を視覚的に確認することができます。既存のワークフローツールとは異なり、ユーザーは SQL ファイル間の複雑な依存関係に注意したり、並列化のための追加の作業を行う必要はありません。

実際のプロジェクトでの AlphaSQL の評価と、Cybozu Lab を卒業後の近況についても発表しました。」

■ 機械学習/自然言語処理に関するソフトウェア開発（担当：中谷 秀洋）

笹田 大翔「差分プライベート深層変換モデルによるテキストプライバシーの保護」

「レビューやツイートのようなユーザ生成テキストは、所属する組織や住所に関する情報等のセンシティブな情報が含まれることがあるため、新規データ事業の創出を目的とした第三者への提供にはプライ

バシ保護加工が必須となります。プライバシー保護加工には個人を識別するデータを除去する匿名化が行われてきましたが、攻撃者の知識を仮定する必要があるため、仮定外の知識を保持する攻撃者にはプライバシー漏洩が生じます。攻撃者の事前知識をプライバシー保護加工に組み込まない差分プライバシーでは、ノイズを付加して出力からのプライバシー漏洩を防ぎます。しかしテキストでは、データ構造上直接ノイズを加えることができません。

そこで、ノイズを加えた勾配で学習した深層変換モデルを介することで、プライバシーの漏洩を防ぎながら有用なテキストの生成を目指します。また学習前に概念関係を保持する固有表現を一般化することによって、ノイズ付加量を低減して性能の劣化を抑制します。」

江畑 拓哉「新進気鋭の深層学習モデル、Flow-based モデルで日本語口調変換を試みる」

「近年注目を浴びている Flow-based モデルと呼ばれる画像生成モデルを応用して、和文の口調変換を試みよう、という研究です。サイボウズ・ラボユースでは、Flow-based モデルを作成するためのライブラリ、TFGENZOO の開発を行わせていただきました。」

■ インフラソフトウェア開発（担当：星野 喬）

舛村 康成「高競合状態における並行性制御法最適化の分析」

「本研究では、近代の並行性制御法の中でも特に優れた性能を持つとされる Silo について、その弱点であると考えられる高競合状態の最適化についての研究を行いました。

高競合状態における優れた最適化として知られる adaptive backoff を詳細に分析し、その性能向上の仕組みについて新たな見解を発見しました。

また、それらの分析を踏まえ、adaptive backoff を超える性能を実現する可能性のある、新たな最適化手法の提案を行います。」

山本 祥平「DBMS を実装しながら学ぶ」

「今回の活動では、DBMS (Database Management System) の構造やアルゴリズムを学習しながら、「Database Design and Implementation」(<http://www.cs.bc.edu/~sciore/simpledb/>) で紹介されているプログラムを参考に、C 言語で実装を行いました。トランザクション技術に興味があり、その内のクラッシュリカバリの機能を実装することができました。また、コードレビューを通して、大きなプログラムを書く上で考えなければならないことを学ぶことができました。例えば、関数の適用範囲を明確にしておくことで、実装後のプログラム管理が行いやすくなることです。」

和田 智優「分散データシステムにおける非決定的性能バグの正確な解析」

「クラウドシステムには、クラウドシステム全体のパフォーマンスが低下する PCbug と呼ばれる非決定的性能バグが存在します。

これを自動的に検知する既存手法 PCatch では、検査対象のプログラムに対し手動で用意された入力を用いて検査対象のプログラムを実行することで実行トレースから PCbug を検知します。本研究では、

PCbug を発生しうる入力を自動的に生成します。

提案手法で生成した入力を用いて計装されたプログラムを実行し、得られた実行トレースから、より網羅的に PCbug を検知します。」

■ C/C++によるソフトウェア開発（担当：光成 滋生）

廣江 彩乃「セキュリティログの解析を助ける視覚化ツール」

「不正アクセスを検知するためのセキュリティログは、サーバに溜め込まれています。しかし、文字列として記録される膨大なログにどのような情報が記録されているかを解析することは、解析ツールなどを使わない限り困難です。

そこで、サーバに記録されるセキュアログの情報を、可視化するためのツールを作成しました。可視化することで、背景知識がない人でもログの内容を理解できるよう、支援することができます。私自身が情報セキュリティの分野に興味を持ったきっかけがまさに、不正アクセスを地図上にマッピングしたものであるため、可視化はとても重要であると考えています。

そして、自分の手でそのようなツールを作成したいと考えていたため、本ツールを作成しました。本ツールは、CentOS の /var/log/secure 配下に定期的に保存されるログを対象としています。ツール作成に用いたログが保存されていたサーバは、web サーバなどではないため、ssh アクセス試行のログが保存されるものです。」

木下 和巳「signal システムコールによるホスト型ハイパーバイザー開発調査」

「本研究では、私が SecHack365 で開発したはりぼて OS が動くエミュレーターである EEMU を拡張し、Linux kernel の機能である sigaction システムコールなどを用いて、ホスト型ハイパーバイザーを実装できるかどうかを実装しながら検証した結果について発表しました。実装したソフトの名前は、ghyper です。(https://github.com/kazuminn/ghyper)

また、試行錯誤を繰り返し実装した過程についても発表しました。」

美濃地 正貴「実行ファイルにおけるバイナリ差分アルゴリズム」

「本発表では、実行ファイルにおけるバイナリ差分を構成する手法を紹介しました。edit distance などの基本的な手法を概観した上で、特定のアーキテクチャに依存することで差分を小さくするアプローチを取る Chromium の Courgette について触れました。」

細谷 啓「RISC-V CPU 自作本の実装と執筆」

「CPU を自作する事はコンピュータアーキテクチャを理解するのに最適な手段の一つです。しかし、CPU 自作について書かれた本はそう多くなく、教育機関以外で独学で行うには未だに大変な困難があると言わざるを得ません。特に、ロードストアユニットの高速化やアウト・オブ・オーダー等の高速化技術について実装まで網羅した書籍というのは無いように見受けられます。

そこで本テーマでは 64-bit RISC-V を対象 ISA とし Verilog, Chisel というハードウェア記述言語を使っ



て、愚直な実装のアーキテクチャからパイプラインアーキテクチャなどへと、CPU として機能する状態で段階的に CPU を設計して拡張していくという本を執筆し、リファレンス実装を与えています。発表ではこの段階的拡張をしていく書籍の今や、執筆・実装時に使った CPU デバッグテクニックの話などをお伝えしました。』

■今後の展望

サイボウズ・ラボユースでは、今後も引き続き活動を継続し、社会にとって有用な研究開発を行う若手エンジニアを支援し、新しい世代の技術者の活躍と社会への新しい価値の創造に貢献してまいります。

■サイボウズ・ラボ／サイボウズ・ラボユースについて

公式サイト：<https://labs.cybozu.co.jp/>

サイボウズ・ラボ株式会社は 2005 年に設立され、サイボウズグループの研究開発部門として、次世代の製品・サービスの基盤となる技術の中長期視点で研究開発しています。2011 年 3 月より、若手エンジニアの育成を目的とした育成プログラム「サイボウズ・ラボユース」を運営し、広く社会に才能溢れる多様な人材を輩出しています。

※商標・著作権表示に関する注記については、こちらをご参照下さい。

<https://cybozu.co.jp/logotypes/other-trademark/>