

ボイスチェンジャーサービスの研究開発を通じた インターンによる次世代人材育成

Developing next-generation creators through research and development in voice changer services

白井 暁彦 SHIRAI Akihiko

GREE VR Studio Laboratory, REALITY株式会社
GREE VR Studio Laboratory, REALITY Inc.

中野 友介 NAKANO Tomosuke

GREE VR Studio Laboratory, REALITY株式会社 / 明治大学
GREE VR Studio Laboratory, REALITY Inc. / Meiji University

堀部 貴紀 HORIBE Takanori

GREE VR Studio Laboratory, REALITY株式会社 / 明治大学
GREE VR Studio Laboratory, REALITY Inc. / Meiji University

民間企業における研究開発において「ボイスチェンジャーサービスの研究開発」をテーマに、学部学生の長期インターンを実施した。一般的に知られている短期インターンおよび採用プロセスでは、どうしても「出来上がった優秀な学生を選ぶ」という視点になりがちであるが、社員と同等のリモートワーク環境を提供し、単なる技術習得やアシスタント的な作業ではなく、研究開発を自発的に推進するための基礎力を学生の視点で実務を通して獲得させながら、学術活動としての論文執筆や外部発表、社内での展開などを行った。学術研究の社会実装を人材育成と研究開発の両輪で行う、次世代人材育成の方法論として共有する。

キーワード：長期インターンシップ、転声こえうらない、音声信号処理、人材育成、産学連携

1. はじめに

産学連携は大学における研究教育活動、ひいては社会実装や継続的な産業の発展において重要な要素である。我が国の産学連携活動の中で、学生のインターンシップをその連携に組み込み「イノベーションサイクル」として活用する方法論の共有はそれほど多くない。海外、特に米国・英国など英語圏インターンシップについては様々な事例が報告されているが^[1]、多くの場合は個々の企業の採用活動を中心とした視点でのレポートであり、デジタルコンテンツ分野における具体的な事例にかかわる報告はあっても^[2]、多くは個々の企業におけるノウハウであり、オープンな形で共有されない。特に、学術的な貢献や高等教育がなすべきエコシステム、ひいては関わる学生の主体的な成長について注目した設計や方法論についての学術文献はほとんど存在しない。我が国においても近年、インターンシップの必要性は文部科学省にも認識されており、学生の就職活動の中で存在感を増している^{[3][4]}。

現状、日本で行われる多くのインターンシップは数日～2週間程度の短期プログラムであり、企業の活動紹介とハッカソンなどのグループワークにとどまるケースは少なくない。2020年以降のコロナ禍において、このような物理的に参集する場を作ることは難しくなった。海外企業へのインターンシップも、留学以上に難しい。一方、リモートオフィスの普及により、首都圏外の学生にも参加する機会が作りやすくなるなど環境の変化は進んでいる。短期インターンシップは学部生向けの就職活動機会としての一定の認知や効果を見せているが、学生側の視点では、志望する企業による採用審査という向きもあり、純粹な成長の場とは考えづらいだろう。より高度なプロフェッショナルの育成という視点でも、大学院を含めた高等教育での研究教育活動や、その研究成果の社会実装と関連がある企業活動に学生が関わることで、企業や大学のイノベーションサイクルに、人材育成という媒体が加わり、継続的な産業の発展を狙う活動につながる。欧州の高等教育では既にこのような産学連携が単位・学位認定システムとともに整備されている^[5]。筆者(白井)自身はフランスの高等教育機関にてポスドクを経験しており、学生インターンシップを通じた産学連携の実践とイノベーションサイクルを目の当たりにしてきたが^[6]、一方で

制度的に整備された環境においては、その個別の取組における工夫や、学生視点の気づきが方法論化しづらいことも認識している。日本におけるデジタルコンテンツ分野の次世代技術開発を行う人材育成に貢献すべく、産学連携、そして学生の視点で方法論を具体的な研究開発の事例を通してまとめていきたい。

2. 技術分野

本稿で紹介する事例の技術分野として、GREE VR Studio Laboratoryが開発・運用しているボイスチェンジャーサービス「転声こえうらない」をはじめとする音声信号処理である。グリーグループでメタバース事業を推進するREALITY株式会社にある研究開発部門であり、研究開発としては音声信号処理だけではなく、アバターSNSにおけるサービス開発、サーバ技術、機械学習、モバイルやサーバプラットフォームにおけるリアルタイム処理などを扱っている。また本論文はコロナ禍におけるリモート環境でのインターン手法、研究活動のDX化なども含まれる。

3. 経緯と成果

インターンの堀部貴紀(当時・学部2年生、以後TH)は、アルバイト社員として2019年10月に参加した。既に2019年7月にリリースした実験的ボイスチェンジャーサービス「転声こえうらない」のエンジンとして採用した音声分析合成システム「WORLD」(<https://github.com/mmorise/World>)の開発者である森勢将雅(明治大学准教授)より紹介され、今後の研究課題を学びながら解決する「研究開発アシスタント」として採用した。

初期段階ではリサーチよりもエンジニアリングの習得に比重を置き、Google Cloud Platform (GCP)上でのPythonによる「転声こえうらない」の利用者データの分析を担当させた。ローカルのPython環境ではなく、あえてGCPを利用する理由は(1)「転声こえうらない」のサーバシステムがGCPで構築されていた、(2)クラウドによるマイクロサービスを構築したい、(3)大学で教える機会がない、という3点が大きい。(2)は時給制のアルバイトであることも関係している。バッチ処理で実行される処理はリアルタイム処理に比べて結果が出やす

く、作業単位を分割しやすい。一方で、本番運用した場合の不具合は社員が対応する必要があるため、品質を維持するためには十分なレビュー体制が必要となる。サービス利用者の選択した選択肢やユーザの音声分析(平均基本周波数・信号対雑音比)、TwitterでShareされたかどうかといった細かな分析のためのツール開発を通して、Pythonだけでなく、GitHubやSlackを使ったコーディングコミュニケーションや、ユーザデータを扱う上での守秘、セキュリティなどを社内のe-Learningシステムや、インフラ部門との交流を通して学ぶ機会を作り出した。外部公開されているサービスとしてはTwitter@koeuranaiにおいて、前日のサービス使用状況をツイートするbotとして運用されている。午前9時15分前日の利用者データの分析を開始し、分析結果のうち人気ランキングを投稿、18時にShareされた作品の紹介をするデイリーバッチ処理であるが、基礎的なスキルの習得と実際の運用を通した予想外の不具合などを学びながら、本番運用に影響がない貢献と経験を積み上げるステップとなった。なお並列して、Google APIでのSpeech-to-Textやpyworldなども使用し、音響学会SIGMUSに向けた論文投稿も指導を行なった^[7]。2020年3月にはコロナ禍によるリモートオフィス化という環境の変化もあったが、初の学会発表にて学生優秀発表賞を受賞した^[7]。なお本人は「オンラインではない学会を知ってはいるが、発表としてはオンライン以外の学会を知らない世代になった」と語った事を特筆したい。

2020年6月、THには次のステップとして「WORLDをiOSに移植して高速化するプロジェクト」に移行することとした。WORLDはC++で書かれたオープンソースプロジェクトであり、iOSプラットフォームの新しい標準であるSwiftをベースとしたプロジェクトにブリッジさせるという開発寄りの作業であり、コンパイラの知識やSwiftの知識が必要となる。これも「大学では学ばない技術」を意識したアサインである。C++を学ぶ大学学部教育は存在するが、Swiftという比較的新しい言語をおしえる授業や演習は多くはない。またモバイルプラットフォームでの音声処理、独自のクラス設計、コンパイラやリンカの働き、デバッガーやパフォーマンス分析、音声信号処理はFFTの高速化など学ぶ要素は多い。週3日程度、10ヶ月ほどの期間の開発を経て、移植及び高速化は成功した。成果についてはREALITY社エンジニアブログ^[9]および、音響学会に寄稿している^[10]。

さらに2020年12月にはTHの後輩である中野友介(学部2年、以後TN)の受け入れも行なった。初期段階ではプログラミングの素養をはかるため、ウェブフロントエンドの学習からはじめた。JavaScriptとVue.jsのベンチマーク比較、Flutter、ドキュメント翻訳、Pythonでの機械学習のための調教データの準備、Google Analyticsを使ったデータ収集の改良、具体的にはAnalytics.jsを最新のGoogle Tag Manager (gtag.js)に移行させる作業を通して「誰でもできること」を品質高く実施する習慣、レポーティングスキル、コミュニケーション習慣などを中心に基礎固めを6ヶ月ほど行なった。その後は、Unityにおけるリアルタイム音声信号処理およびインタラクティブ技術の開発をアサインしているが、入社当時よりも新しい言語に対する自律学習および習得速度が上がっており、音声分析による動的アニメ生成など多分野に横断する才能が開花している。

4. まとめ：今後に向けて

日々の作業スタイルは、朝9:30開始、夕方18:30までが定時であり、朝会(Zoom)にてその日の行動方針や目標などを各インターンが自律的に計画し、Google Slidesにて1日1枚の日報を書き、夕会にて成果を発表する。必要であれば個別に集中作業としてミーティングやペアプログラミング、もしくはモブプログラミング(一人のコーディングを複数人が見ながら学習する)を行うこともある。ラボ内で常に他のメンバーのプロジェクトを見ながら日々進行するため、外部向けのコンテンツやPoCプロジェクト、学会発表などがある場合はQA作業やレビューなどに積極参加させる。大学の研究室と異なる点は時給制のアルバイト社員であり、時間に対するパフォーマンスは常に評価される点である。終了時間は特に許可がない限りは残業しない方針で運営している。学生視点での特色として、グリーグループはアルバ

イトであっても社員と同じ扱いをするため、社員の全社ミーティングや総会といった比較的機密レベルが高い会議にも出席し、メタバース開発の現場のサービス開発の現場を学ぶ機会がある。また社外向けのBlog執筆や特許出願も推奨されており、発明報奨金などの制度も社員と同様に提供されている。なおコロナウイルスに対するワクチン接種も職域接種で実施した。

短期インターンおよび採用プロセスでは、どうしても「出来上がった優秀な学生を選ぶ」という視点になりがちであるが、近年の大学の環境を顧みると「人を育てる」という視点で、回り道をしながらも、できるだけ柔軟に幅広い技術や表現力を、自分の中に「垂直に」醸成できる人物を育成していくことが近道にも思える。もちろんそのためには相当の時間も必要であるが、研究開発を自発的に推進するための基礎力を学生の視点で実務を通して獲得させながら、学術活動としての論文執筆や外部発表、社内での展開など、企業活動を通した社会実装スキルを高めていく両輪の視点の方法論は重要であり、今後も長期インターンシップ事例の共有機会があればまとめていきたい。

謝辞

本論文を執筆するにあたり、GREE VR Studio Laboratoryのインターン各位、REALITY株式会社・荒木英士氏、および明治大学・森勢将雅先生に感謝を記します。

参考文献

- [1] リクルートワークス研究所: "海外におけるインターンシップ最新事情" https://www.works-i.com/research/others/item/160322_Internship_jp.pdf (Accessed 2021-07-31).
- [2] デジタルハリウッド株式会社, PR TIMES: "夏期3日間でおこなう実務体験型インターンシップ『DHW Boat Config』を開催", <https://prtimes.jp/main/html/rd/p/000001832.000000496.html> (Accessed 2021-07-31).
- [3] 文部科学省: "大学等におけるインターンシップの推進", https://www.mext.go.jp/a_menu/koutou/sangaku2/1346604.htm (Accessed 2021-07-31).
- [4] 文部科学省: "産学連携による実践型人材育成事業－長期インターンシップ・プログラム開発－(平成18年度採択分)の最終評価について", https://www.mext.go.jp/a_menu/koutou/sangaku/haken/1319022.htm (Accessed 2021-07-31).
- [5] 富山哲男: 「ヨーロッパにおける産学連携」精密工学会誌, Vol.71, No.1, 2005.
- [6] Patrick Corsi, Hervé Christofol, Simon Richir, Henri Samier: "Innovation Engineering: The Power of Intangible Networks", John Wiley & Sons (2013).
- [7] 堀部貴紀, 石原達馬, 白井暁彦, 森勢将雅: 「『転声こえうらない』利用者の基本周波数分析」情報処理学会音楽情報科学研究報告(2020年), vol.2020-MUS-127, no.18, pp.1-
- [8] 堀部貴紀, 白井暁彦, 森勢将雅: 「『転声こえうらない』を通じたボイスチェンジャー品質改善のための定性分析と考察」日本音響学会2021年春季研究発表会(2021年).
- [9] REALITY株式会社 note記事: "iOS Accelerateでボイスチェン高速化に挑戦したインターン学生の話", https://note.com/reality_eng/n/0cd9bd157df3 (Accessed 2021-07-31).
- [10] 堀部貴紀, 橋本順之, 白井暁彦, 森勢将雅: 「FFTライブラリを対象とした実時間Vocoderの速度比較」日本音響学会2021年秋季研究発表会(2021年).
- [11] 白井暁彦, 堀部貴紀, プレディキナー リュドミラ, 山崎勇祐, 坂口塔也, 中村清人, 栢之間諒汰: 「【研究ノート】Mozilla Hubsを用いたオンラインイベントのWebVR化」, DHU JOURNAL Vol.07 2020. (2020/11/26)