
研究ノート

法科大学院教育における AI 音声認識技術を使った収録システムを導入する際の予備的検討

富 崎 おり江

要旨

名古屋大学法科大学院では、限られた期間内に新司法試験を受験できるレベルに引き上げることが求められており、そのためには効果的に教育することが課題となっている。法学教育の初期段階において例えば、聞き逃したところや理解できなかったところを随時確認できるようにすれば学生はより理解を深めることができると考えた。名古屋大学法科大学院は講義を繰り返し視聴できるインデックス付き講義収録システムを独自開発した。2005年から講義を収録し基本的知識を定着させる環境を提供してきた。

近年、講義収録システムは講義の動画を配信するだけでなく、他に学修につながる付加データを付けて提供する試みがなされている。例えば、動画から得られる音声、テキスト化し、人工知能で解析し、重要な単語を自動で抽出する。収録した講義に重要な箇所を自動でタグを挿入する。授業後、学生はスマートフォンやタブレット端末で講義の重要箇所をタグから再生することで効率よく復習を行うといったことである。教員からすれば「あたりまえ」を学生はきちんと理解できているのか、学生からすれば「なんとなく」しか理解していなかったことを吸収できているのか可視化する。似たような試みは初等教育で既に始まっている¹⁾。

本研究では法学の基礎知識の復習・理解度アップのため収録データに付加データを付けて提供することは法曹を養成する法科大学院にも有用となると仮定した。まずは AI 音声認識技術を使った収録システムを導入する

1) 黒板メーカーのサカワは、音声認識技術と人工知能 (AI) を組み合わせた授業支援のための新たなシステム「Josyu (ジョシュ)」を開発 <https://ai-josyu.com/>

際の予備的検討を行った。模擬授業のアンケート結果から、音声認識の精度を上げる必要があることがわかった。コーパスの作成、収録の際の問題、収録機器の経費といった問題が浮上した。

本稿では実際の現場で様々な問題が渦巻いている現実を紹介する。多様なメディアを高度に利用した授業の発展に参考にされたい。

1 はじめに

現在 ICT の法科大学院教育での活用・普及は政策上の課題になっている。平成 25 年 9 月に法曹養成制度改革推進会議を設置し、平成 27 年 6 月 30 日に「法曹養成制度改革の更なる推進について」という決定を公表した。この決定の中に「法曹を志す者の誰もが法科大学院で学ぶことができるよう、ICT（情報通信技術）を活用した授業の実施状況などの事情を適切に考慮するものとする」とされているのである²⁾。

平成 30 年の中央教育審議会大学分科会法科大学院等特別委員会では、法科大学院等の抜本的な教育の改善・充実に向けた基本的な方向性として、〈方向性①〉法科大学院と法学部等との連携強化—法学部に「法曹コース（仮称）の設置を奨励し、法学部が法科大学院と連携して体系的・一貫的な教育過程を編成することにより、法曹志望が明確な学生等に対して学部段階から効果的な教育を行う。〈方向性②〉法学未修者教育の質の改善—進級に当たっての質保証プロセスを導入するとともにきめ細やかな指導を効果的に行うことが提示された³⁾。これらを遂行するには ICT を活用した教育支援が必要である。

大学における多様なメディアを高度に利用した授業は政策上の課題である。中央教育審議会においては、高等教育機関としての質の担保と向上、そして革新のためには、ICT を活用した教育を進めていく必要があると述べている⁴⁾。

2) 法曹養成制度改革推進会議とその関係組織については http://www.kantei.go.jp/jp/singi/hoso_kaikaku/ を、その決定は http://www.kantei.go.jp/jp/singi/hoso_kaikaku/pdf/honbun.pdf を参照

3) https://www.mext.go.jp/b_menu/shingi/chukyo/chukyo4/041/siryo/_icsFiles/afieldfile/2018/03/14/1402456_002_1.pdf を参照

4) 中央教育審議会第 18 回大学分科会制度・教育改革ワーキンググループ (http://www.mext.go.jp/b_menu/shingi/chukyo/chukyo4/043/siryo/1409011.htm 資料 6 大学における多様なメディアを高度に利用した授業についてを参照)

このような流れを踏まえ全国的に法科大学院と法学部の連携促進による法曹養成を目指さなければならない。学生の時間的・経済的負担の軽減や法曹養成制度の信頼性・安定性の質保証を図ることが求められている。これらを実現するための手段の一つとして、ICTを活用した教育支援環境を提供してきた経緯があり、本研究の着想に至った。

2005年から講義収録システムを利用している。このシステムは講義を聴きながらもう一度聴きたいと思った時点で Word ファイルにインデックスを付け、講義終了後、インデックスをクリックすると聴きたいと思ったところから再生できる機能がついている。実際に導入してみると、学生は講義を聴きながらノートを取るのが精一杯で、インデックスを付ける余裕は無く、収録データの再生のみの利用となっている。注目したのは講義収録システムの活用方法として、講義の文字起こしをすることで復習する学生がいることである。この学習方法のよし悪しは別として、初学者は法律用語が聞き慣れないのが要因となっていると推測した。

第三次人工知能ブームによるディープラーニングの登場やスマートフォンへの搭載により、急速に普及した AI（人工知能）の技術に「音声認識」がある。この技術を使った講義収録システムが登場した。収録したデータをテキスト化する。講義の収録データを再生する際に字幕の表示やキーワード検索でもう一度聴きたいと思った部分をピンポイントで再生できる AI 音声認識講義収録システムを法科大学院で利用するための予備的検討を行った。

今後、大学における多様なメディアを高度に利用して行われる授業が増えていく。本稿では、これから抱える課題を明らかにし、今後の多様なメディア授業の発展を目指す。

2 方法

2.1 調査時期と調査対象者

AI 音声認識講義収録システムを法科大学院で利用するための予備的検討を行うため、模擬授業を行った。

模擬授業は 2019 年 2 月 6 日であった。内容は定期試験の解説であり、

その様子を収録した。参加者は未修1年生を対象に募集したところ、調査対象は7名になった。模擬授業の受講生として質疑応答にも参加してもらった。

視聴実験は2019年2月20日であった。模擬授業の様子をスマートフォン、タブレット端末およびノートパソコンで視聴し、その使い勝手についての感想を質問紙によるアンケート形式で調査した。

システムはDaAlps for Education⁵⁾を利用した。

2. 2 調査内容と結果

質問事項は次のようなものであった。

- ・システムの検索画面、タグ表示画面、全文表示画面の操作性に関する問い
- ・現在利用している講義収録システムの利用頻度とタイミング
- ・DaAlps for Education が利用できるようになった場合の利用頻度と利用したい場所

本稿に関係する質問と回答を下記に示す。

「授業の内容が全文テキスト化されることについて便利だと思いますか？」の質問に対し全員が「便利だと思う」という評価だった。そう思える理由の自由記載欄に書かれた回答を示す。

- ・動画を再生しなくてもテキストを見れば、授業の大まかな流れが思い出せる。
- ・誤字は多いものの、検索しやすい。
- ・知りたいところにすぐにアクセスできるから。
- ・先生が何の話をしていたか見返しやすいから。ただ、誤字が多く句読点が無いため読みにくい。
- ・聞き落としや先生に質問するまでもない質問・疑問を解消できるから。

「タグキーワードから再生できる機能は便利だと思いますか？」に対する評価と自由記載欄に書かれた回答を示す。

5) 共信コミュニケーションズ株式会社開発の講義収録コンテンツソリューション

「思う」3名

- ・授業のどの部分を再生したかったのか見つけ出しやすくなる。
- ・90分の授業を、再度時間をかけて視聴するのは面倒。文字だと、本当ならもっと時間がかかるのを5分程度で終わるので便利だと思う。
- ・講義全体を通してどんな話をしていたかわかりやすいから。

「どちらも言えない」2名

- ・検索機能自体は便利だけれど、可能であれば該当箇所に飛んで欲しい。
- ・検索結果から、その場面の動画に用語をクリックすることで再生してほしい。

「思わない」2名

- ・キーワードの該当箇所に飛ばない。
- ・ブラウザのCtrl + Fを使わなければならない、2度手間になる。

「思う」と答えた学生は、90分の授業をシークバーで視聴したい部分を探さなくても、全文テキストやタグキーワードにリンクされているタイムコードにジャンプして再生できるため、時間の短縮になると考えていることがわかる。「どちらも言えない」や「思わない」と答えた学生は、検索したキーワードからピンポイントでちょうど思い通りの該当箇所から再生できなかったためこのような評価をつけたと推察できる。これはタイムコードの区切り方によるもので、区切り位置によっては該当箇所よりも前の部分から再生されることは想定範囲内である。

実験の参加者全員が「便利だと思う」という評価だったため、AI音声認識講義収録システムを法科大学院で提案する誘因になった。ただ、自由記載欄の「誤字が多い」という回答が気になった。音声認識の精度を上げるため何ができるか検討したところ、法律用語コーパスの作成、収録の際の問題、収録機器の経費といった問題が浮上した。

3 法律用語コーパスの作成作業

AI音声認識講義収録システムを利用したアンケート結果によると、前述したようにテキスト化したデータに誤字が多いという回答があった。収録した講義やテキスト化したデータは受講生のみが利用可能とする運用が前提である。学生が復習で利用する際は教科書や講義ノートを見ながら収

録した講義データを再生する組み合わせで利用することを想定しているため、言葉の齟齬が多少あっても利用に耐えられる。しかし、法律の専門用語を正しくテキスト化して欲しいというのが理想である。微妙な表現の違いで真偽が左右され得るため、文字起こしがより正確にならないと使いづらい訳である。それならば収録したデータを聴きながら手作業でテキスト化したデータを修正すればいいのだが、90分の講義を週15コマ収録しているため、この方法は現実的ではない。

今回の実験で利用した音声認識エンジンは Google Cloud Speech-to-Text である。汎用性が高いものの法律の専門用語をテキスト化するのは難しかった。そこで、音声認識エンジンを法律専門用語にチューニングできないか検討することにした。

まず、法律用語のコーパスをどのように作るかの課題がある。言語モデル⁶⁾を作るのに5メガ単語必要と言われている。課題となるのは、これらの作業にどれほどのコストが必要になるかである。筆者が、音声認識エンジンのチューニング作業の委託に関わったことがある。模擬裁判や模擬相談を音声認識で文字起こしするためである。教材等のテキストデータと模擬裁判および模擬相談の収録データを提供し作業をお願いしたところ、150万円ほどかかった。

法律用語のコーパスを作成するには音声データとテキストデータが必要になる。テキストデータは授業のレジメや司法試験の過去問題をテキスト化したものを利用することは可能である。問題は収録データである。もともとコーパスを作成する目的で収録していないため、音が鮮明ではなく、雑音が多いため、法律用語のコーパス作成には不向きであった。

AI 音声認識技術のためのコーパスをいかに生成するかは作業量や費用面もそれなりにかかる。研究対象になるぐらい難しいテーマと言われているが理由がよくわかった。

6) 膨大な量のデータから単語のつながりを予測判定し、より正確な文章を組み立てる。あらかじめ蓄積したデータから使用する単語の出現率を算出し、単語を文章化すること。

4 収録する際の問題点

AI 音声認識講義収録システムで良質なテキストデータを生成するためには、雑音を少なくし声をクリアに録音する必要がある。雑音が多く、収録した講義の声が不鮮明な音声だと、AI が正しく認識できず、逆に講義と関係のない音をテキスト化してしまうなど、質の良いテキストをアウトプットできない。AI 音声認識講義収録システムを最大限に発揮するためには「環境」と「マイク」の配慮が必要である。「環境」としては、私語やその他の雑音の無い状態が理想である。講義中に収録するため、私語を慎んでもらうのは当然として、ノートを取るためにキーボードの入力音や紙をめくるといった雑音を立てないように注意喚起が必要である。「マイク」についてであるが、講義室で利用している拡声マイクの場合、周囲の音を広く拾ってしまうため特定方向の音だけを拾う指向性のマイクを使うのが理想である。ただし、マイクの入力レベルやマイクの位置などの調整が必要になる。法律基本科目のうち 11 科目を週 15 コマ収録しなければならないが、収録の度にマイクの位置や入力レベルに気を配るのは負担になる。これでは本来の目的の教育支援のためのシステム利用から逸脱してしまう。

話し方も大切なポイントである。音声認識のプレゼンテーションに参加した際、発表者が音声入力すると正確に認識されるのに対し、筆者が音声入力すると音声認識精度が下がった。話し方のコツが必要なのである。音声認識の精度を上げるため、90 分の講義中ずっと活舌よくアナウンサーのような発音を続けるのは難しい。

5 収録機器の経費

講義を収録するにはマイク、オーディオミキサーアンプなどの音響設備、カメラ、アップコンバーターなどの映像設備および収録した音や映像をサーバに転送するための収録設備が必要である。2 教室分で 1400 万円ほどかけて整備した。それなりに数年経過すれば機器の故障が発生するようになるため機器の更新が必要になる。現時点で数千円から数万円の価格のマイクを内蔵した防犯カメラやビデオ会議用カメラがある。このような手ごろな価格のカメラやマイクを活用し講義収録システムに組み込めないの

か意見が寄せられるが、そう簡単にはいかない。以下に理由をまとめる。

5. 1 安定したフレーム数が保証されない

通常のカメラ・ビデオ映像は、1秒間に30コマで構成されている。Webカメラのように撮影した映像をネットワーク経由で送る場合、ネットワークの混み具合などで、十分な転送速度が得られない場合にはコマ飛びが起きる可能性がある。また、システムにもよるが、コマ飛びが起きると映像と音声のタイミングズレが生じることがある。講義の収録で映像と音声がずれると内容に集中できなくなる問題がある。

5. 2 防犯カメラシステムは外部からの制御が難しい

通常の防犯システムは、基本的に録画し続けるシステムとしてパッケージ化されているため、外部からの録画開始・停止操作や収録後の動画データの外部への取り出しが極めて難しい。通常、防犯カメラには録画開始・停止という概念がなく、何時から何時までの録画という時間設定になるため、その都度の任意の操作に困難を極める。また、安価な防犯システムの動画フォーマットは、そのメーカーの独自フォーマットである場合が多く、Windows上での標準再生ができない場合が多いことも課題である。

5. 3 セキュリティーに対する不安

ネットワークに接続される安価な防犯システムの場合、収録した動画はメーカーサイトから提供される再生用アプリを利用し再生可能となる。よって、メーカーはすべての収録機が存在が把握できており、再生機側（一般のPC側）も把握が可能になる。これはそのシステムに対するハッキングも簡単にできる状態になるため、収録した動画が流出しかねず、セキュリティー上の危険がある。

6 まとめと今後の課題

高等教育機関としての質の担保と向上、そして革新のためには、ICT を活用した教育を進めていく必要がある。その方法の一つとして AI 音声認識技術を使った収録システムの予備的検討を行った。本研究から十分な財源が確保できない点とノウハウ不足が明らかになった。

地方も大都市と同等の教育を提供するには、人材の不足や十分な財源が確保できないという問題がある。地方の大学等が連携して対応するためには、ICT を利用した教育は必ず必要になる。大学における多様なメディアを高度に利用した授業は政策上の課題である。ICT 教育を進める仕組みを作るなら、より地方の高等教育の革新へとつながると期待したいが、これには資金と技術支援の確保が肝要である。

付記

本研究は公益財団法人日東学術振興財団の助成を受けました。また、本研究にご協力いただいた共信コミュニケーションズ株式会社提供の講義収録システム「DaAlps」開発者の皆様および模擬授業の収録と視聴実験にご協力いただいた皆様に深謝申し上げます。

