

生成 AI を利用した書籍作成プロセスの開発

大石耕生[†] 大場 みち子[‡]京都橋大学 工学部情報工学科[†]

1 研究背景

近年、IT は極めて速いペースで進化している。特に、人工知能、機械学習、データサイエンス、ブロックチェーンなどの分野では、新しい発見や技術革新が連日のように報じられている。このような IT トrendは社会に大きな影響を与え、一般消費者にもこのスピードに追従することが求められている。

この状況において、IT のトレンドトピックを扱う書籍は、最新の情報を提供するための重要な媒体となる。しかし、従来の書籍出版プロセスは、著者の原稿作成から編集、校正、印刷、配布まで数ヶ月から一年以上もかかることが一般的である。この長い時間枠では、書籍が市場に出る頃には、含まれる情報が古くなってしまい、読者の手に渡る時点で既に陳腐化している可能性が高い。このギャップを埋めるためには、出版のプロセス自体を根本から見直し、より迅速に対応できるように改革する必要がある。以上より、トレンドトピックの書籍化では、原稿の作成から出版までの時間を大幅に短縮する新しいプロセスの開発が急務である。そこで、生成 AI による物語の生成や映画のシナリオの生成など創造性を支援する取り組みに着目する [1][2]。

2 研究目的

本研究の目的は、トレンドトピックに関する書籍の作成時間とコストを大幅に削減する新しいプロセスを開発することである。

3 課題と解決アプローチ

本研究で取り組む課題は2つある。

課題1 トrendトピックの書籍出版プロセスの効率化

現在の出版プロセスは時間がかかりすぎ、トrendトピックに関する書籍が市場に出る時点で内容が既に古くなってしまう可能性がある。このギャップは、情報の即時性が求められる現

代において、特に技術や流行が日々進化する分野で問題となる。

課題2 手作業の削減

編集および校正のプロセスでは、手作業が多く、多大な時間と労力が消費される。手作業が多い現在のプロセスは非効率であり、コストと時間の削減が可能な自動化および最適化が必要である。

4 トrendトピックの書籍出版プロセスの効率化

最新のトrendトピックに関するテーマでは多数のシンポジウムやセミナーが開催されている。そこで、シンポジウムやセミナーの企画、資料、およびプレゼンテーションを活用する。手作業で行っていた従来のプロセスを生成 AI 等を利用して、音声からテキストへの変換を含む作業を通じて、出版プロセスを合理化する。

最新のトrendに関するシンポジウム資料を使用して、専門家の最新の知見や研究成果を書籍の原稿作成に活用する。これにより、即時性と専門性を確保し、トrendに沿った内容を提供できる。図1は新しい書籍出版プロセスの例である。

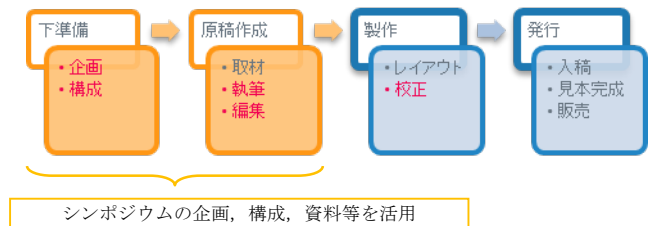


図1 新しい書籍出版プロセスの例

次に、シンポジウムのプレゼンテーションやディスカッションを録画した動画から、音声ファイルに変換する。このステップで専門家の議論や意見を直接キャプチャし、書籍のコンテンツに反映させるための素材を得る。

変換した音声ファイルは、文字起こしツールを使用してテキストデータに変換する。現在の文字起こしツールは高度な音声認識技術であり、

さまざまなアクセントや環境下での音声も正確にテキストに変換できる能力を持っている。これにより、高い精度で素早くテキスト化を行い、人間の編集者が行う作業を効率化する。

文字起こしツールによって生成されたテキストは、書籍としてのフォーマットに合わせて整形する。この段階では、文字起こしツールによるテキスト化の精度はかなり高いが誤字脱字、句読点の不備、文脈の不明瞭さなどが生じるため、生成 AI を利用してこれらを修正する。そして、出来上がったテキストを初期原稿とする。

5 手作業の削減

手作業の削減では、シンポジウムやセミナーで発表した発表者による添削と、そのプロセスの生成 AI による自動化を図る。

次に自動化のアイデアを説明する。

A) 原稿の比較分析

文字起こしツールによるテキスト化と編集プロセスを経て作成された初期原稿は、シンポジウムやセミナーでのプレゼンテーションやディスカッションを基にしている。そのため、発表者に初期原稿を直接添削してもらい、内容の正確性と専門性をさらに向上させ、書籍として適した形に修正する。その後、添削された原稿と初期原稿を比較し、どのような変更が加えられたかを分析する。この分析には、修正された文言、追加された情報、変更された構造などが含まれる。この比較分析から、専門家が重視するポイントや一般的な添削のパターンを把握する。

比較分析の結果を基に、将来の添削プロセスを自動化する。これにより、手作業による添削の時間と労力を大幅に削減し、効率的な編集プロセスを実現する。

B) 発表動画の分析

シンポジウムの動画から発表者の声のトーン、大きさ、強調などを分析し、発表者の主張するポイントを抽出する。MM-VID[5]はマルチモーダルなビデオ分析ツールで、音声や映像の要素を統合して分析することができる。この分析により、発表者の意図や重要なポイントをより正確に把握し、それを原稿の編集に反映させることができる。これらを活用して、初期原稿の編集過程の自動化を目指す。

6 予備実験と結果

シンポジウムの動画からテキストデータに変換し初期原稿を作成する予備実験を実施する。

6.1 実施手順

予備実験は、6名の学生により実施した。シンポジウムのプログラムに基づいて動画を分割

して発表ごとに初期原稿を作成した。実施手順は以下の通りである。

- (1) シンポジウムの録画動画を音声ファイルに変換する。
- (2) 音声ファイルを OpenAI の Wisper[4] を使用してテキストデータに変換する。
- (3) テキストデータの誤字脱字、句読点の挿入など ChatGPT[3] の API を使用してテキストを整形する。

6.2 実験結果

予備実験で実施した発表に対する各作業の処理時間を図2に示す。生成 AI を活用する新しい書籍出版プロセスは、手作業で原稿を作成する従来の方法と比較して、原稿作成の速さで明確な優位性を示すことができた。

発表に対する各作業	処理時間
動画を音声ファイル(wav)に変換	10~15分
Wisperの文字起こし起動時間	6~7分
APIでテキスト整形起動時間	12~13分

図2 発表に対する各作業の処理時間

予備実験で活用したシンポジウムは各発表が20分前後であった。図2から、20分前後の動画から初期原稿を作成するために、必要な処理時間は35分程であった。実験結果より初期原稿作成という点で、時間的効率が大幅に向上した。

7 まとめ

本研究では、最新のトレンドトピックに関する書籍を作成する新しいプロセスを開発し、関連する時間とコストを大幅に削減することを目的とした。書籍出版プロセスの効率化と手作業の削減という2つの主要な課題に対して、生成 AI の技術を利用してアプローチをとることにした。シンポジウムのビデオからテキストデータに変換し、そのテキストを精緻化する新しいプロセスの効果を検証した。編集プロセスのさらなる自動化と検証を進める。

参考文献

- [1] 日笠 敬大, *川村 天, 栗原(2022)「聡創造性の支援のための物語構造を利用したプロットの自動生成」『人工知能学会全国大会論文集』
- [2] *川野 陽慈, 宇都宮 悠輝, 高屋 英知, 山野辺 一記, 栗原 聡(2020)「*川野 陽慈, 宇都宮 悠輝, 高屋 英知, 山野辺 一記, 栗原 聡」『人工知能学会全国大会論文集』
- [3] OpenAI : ChatGPT ([OpenAI](#))
- [4] OpenAI : Wisper ([OpenAI](#))
- [5] Microsoft : MM-VID ([Microsoft - クラウド、コンピューター、アプリ & ゲーム](#))