

補助事業成果報告書

補助事業の名称	視覚障害者・ディスレクシアのための音声を使った読書方法の研究開発（株式会社想隆社）
補助事業の概要	以下に記す

【研究開発の実施内容と成果】

1. 本研究の概要

YourEyes (<https://youreyes.jp>) は、老眼や弱視を含む広義の視覚障害者、さらには識字障害の一種であるディスレクシア症例を持つ人への読書支援サービスである。2021 年 12 月にプレスリリースが株式会社ポニーキャニオンより行われ、2022 年 4 月に一般に公開された。その後、2022 年、ポニーキャニオンから運営会社の変更があったが、弊社想隆社は引き続き開発チームとして本プロジェクトのコア技術の研究開発に関わっている。

R3 年度は、このサービスのクオリティを向上させるべく、以下の 3 つの観点から本システムの改良のための研究調査を行い、改良版プロトタイプ版の制作を行った。

1. 書籍の版面レイアウト解析技術の開発
2. 中間形式の研究
3. ブラウザ版リーダーの開発のための基礎研究

1.1 書籍の版面レイアウト解析技術の開発

OCR された書籍のページ（版面データと呼ぶ）をレイアウト解析し、柱とよばれるページ上部のヘッダー部や、ページ番号であるノンブルを本文の中に入れないように分離する技術が必要となる。新書、文庫といったシリーズになっている書籍では、定型の版面デザインが行われるため、ある ISBN の本が特定の新書・文庫等の提携フォーマットであった場合、その定型情報からレイアウトの概要を得てレイアウト解析に利用する。OCR エンジンの性能だけに頼らず、外部ナレッジを使った OCR 精度の向上を目指すものである。

外部ナレッジには、Amazon や Google ブックスといった大手オンラインストアでも利用されている有料書誌情報データベース「ブックデータベース」（日外アソシエーツ社）を利用する。

1.2 中間形式の研究

中間形式については従来の音訳ボランティアの方法を取り入れ、音訳ボランティアの読みと同等の機能を実現できるように中間形式のフォーマットを策定した。10 余りの独自タグをボランティアツールでも使えるようになり、従来の音訳ボランティアでもスムーズに YourEyes のサービスのボランティアに移行できると考えられる。

1.3 ブラウザ版リーダーの開発のための基礎研究

YourEyes はいわゆるスマホアプリと呼ばれる独自アプリケーション型であるが、視覚障害者がアプリ

をインストールするハードルやスマホ自体を操作するハードルがある。一方、ブラウザはどんなパソコンにも標準搭載されており、視覚障害者にとってもキーボードというなじみのある入力機器を使えるため、ハードルはスマホと比較して低い。ブラウザで機能が完結できることが可能かどうか調査し、ブラウザ版の試作を行った。

2.研究開発の成果

2.1 書籍の版面レイアウト解析技術の開発

AI による画像認識に加えて巨大書誌データベースのナレッジを利用したことで精度が向上した。従来方式と比較して最大で 3 割程度の向上が見られた。また、本システムの評価については、専修大学の植村先生、野口先生に本システムを利用していただき、同様のサービスとの比較も行っていただいた。その結果、マイクロソフト社がアプリ内課金型で配布している Envision AI などと比べても読書に特化した場合においては、YourEyes のほうが文字認識の精度が優れていることが分かった。

精度比較①（誤読率・誤読数）

	YourEyes	Envision AI	Seeing AI	サリバンプ	Lens	よむべえ	
文庫本	0.8% (3.9)	3.2% (15.4)	3.1% (14.7)	1.0% (4.9)	4.5% (21.9)	1.8% (8.9)	
単行本	1.2% (5.7)	2.2% (10.3)	2.7% (12.9)	1.5% (7.3)	3.1% (14.8)	2.3% (10.9)	※（）内は 1ページ平均誤読数

- ・文庫本、単行本どちらもYourEyesとサリバンプが誤読率が少なく読み間違いが少なかった。
- ・誤読例としては「『つ』を『0』」、固有名詞（氏名、地名など）の誤読が多く見られた。特にSeeingAIとLensに多かった。
- ・Lensは文章として形を成していないことが比較的多かった。

精度比較②（ボックス使用）

	YourEyes	Envision AI	Seeing AI	サリバンプ	Lens
ボックス無し	0.8% (3.9)	3.2% (15.4)	3.1% (14.7)	1.0% (4.9)	4.5% (21.9)
ボックスあり	0.9% (4.5)		2.7% (13.1)		4.1% (19.9)



ボックス使用時の反射

※文庫本の場合
※ () 内は1ページ平均誤読数

・ボックスを使用した場合の方が、より精度が増すかと思われたがボックス底部の亚克力板にフラッシュが反射するなどの要因から精度が劣ってしまった。

精度比較③（標準偏差値）

	YourEyes	EnvisionAI	SeeingAI	サリバンプ	Lens	よむべえ
ボックス無し	3.5	3.1	6.7	2.5	10.0	0.8
ボックスあり	0.5		2.1		2.3	

・誤読数のばらつきは「よむべえ」の0.8が最も小さく、安定した読みを行うことがわかる。
・またボックスありの場合ばらつきは減少した。

（2022年2月24日 画像電子学会第50回VMA研究会発表資料より専修大学山崎航・野口武悟・植村八潮（専修大学）らの発表資料による）

2.1 中間形式の研究

新書・文庫程度の文字ものの書籍に関しては中間形式を用いて人間の読みを保持することに成功した。実際にボランティアをYourEyes上で応募し作業してもらった結果、音訳ボランティア経験者からも、機能が足りず音訳作業ではできなかったとの意見は出なかった。

2.3 ブラウザ版リーダーの開発

アプリでは、技術的に実装できないもの以外は機能を付加するはできるが、ブラウザベースの場合、ブラウザに機能実装されていないものを機能搭載することはできない。

ブラウザは主に W3C と呼ばれる規格団体による仕様が決定されていない技術は搭載されず、また W3C の規格にあってもブラウザベンダーの判断によって仕様通りに実装されていないものも少なくない。まずは、これらを調査し、ブラウザ上で SSML と呼ばれる読み上げの規格への対応を調査した。

以下のようになり、まだ SSML へのブラウザの対応が進んでおらず、今後、ブラウザへの SSML 対応が進むことを期待しつつ、それまでの移行期間は、ブラウザが標準でもつ機能をなんらかの形で補完する必要があることが分かった。

ブラウザ	結果	挙動
Chrome	○	
Firefox	×	タグも読み上げられ、制御できず
Edge	×	タグも読み上げられ、制御できず

Table Web Speech API で SSML を含むテキストを読み上げた場合

補完はひとまず置いておき、標準の機能だけで実装し、<https://app.youreyes.jp/Vtool.html> にて公開した。

書籍確認

表示された内容を確認し、「OK」ボタンをクリックして下さい。

本のタイトル	それでも、日本人は「戦争」を選んだ
本の概要	膨大な犠牲と反省を重ねながら、明治以来、四つの対外戦争を戦った日本。指導者、軍人、官僚、そして一般市民はそれぞれに国家の未来を思いなお参戦やむなしの判断を下した。その論理を支えたものは何だったのか。鋭い質疑応答と縦横無尽に繰り出す史料が行き交う中高生への5日間の集中講義を通して、過去の戦争を現実の緊張感のなかで生き、考える日本近現代史。小林秀雄賞受賞。
出版社	新潮社
版	初版

やり直す

OK

Fig ブラウザ版試作品画面



Fig ブラウザ版試作品画面

2.4 採用・受賞

YourEyes は 2021 年サービスを開始したが、ブラウザ版のリリースなど 1 年を通じて活発な活動を行った結果テスト公開に上場企業での採用等もあった。まだ本研究の成果が 100%反映されたわけではないが、YourEyes が研究の成果を取り入れ、実用化レベルに近づきあること実感があった。また、電子出版制作・流通協議会が主催する電流協アワードにおいて特別賞を受賞した。
(https://aebs.or.jp/aebs_award2021.html)

2.5 当事者からの声

YourEyes には多いときは週に十数通のメールが視覚障害者から寄せられ、改善点や意見が寄せられた。2021 年度は、運営会社の変更などの想定していなかった問題があり、視覚障害者をモニターとした意見聴取はあまりできなかったが、次年度の令和 4 年も引き続き行っていく必要があると考えている。

3. 当初目標についての達成度

3.1 レイアウト解析機能

レイアウト解析機能に関しては、ナレッジベースとのハイブリッド化で申し分のない精度が出たと考えている。ただし、YourEyesBox を使った利用的な環境を条件にしている。今後はこの BOX に依存しない撮影環境への対応が必要となることが分かった。

3.2 中間形式の研究

ボランティアの応募数 が運営会社の変更という想定外の問題のため、予想外に伸びなかったため、引き続き音訳ボランティア経験者などの意見を聴取する必要がある。

3.3 ブラウザ版リーダーの開発

ブラウザへの SSML 対応が思ったほど進んでいないことが調査の結果判明し、ブラウザ上で動く技術（たとえば JavaScript）などを使ってさらに音声記号列を音声化する必要がわかった。これについても調査と試作を続行する必要がある。

3.次年度（令和 4 年度に向けての課題）

今年度、令和 3 年度は、データベースのナレッジと AI をハイブリッドにしたレイアウト解析を導入し、精度向上を得たが、レイアウトを新書や文庫程度の文字を中心にしたレイアウトが複雑でないものを使った。サービス利用者からも意見が出たが、複雑なレイアウトや BOX に依存しない撮影環境でもレイアウト解析を行い、文字認識率を高めることが必要であることが分かった。

さらに中間形式もそれに対応した拡張が必要になる。たとえば数式の読みなど専門性の高いものはさらに音訳なども複雑になる。また、著作権保護の観点から、本サービスでは手元に書籍があることをサービスの前提にしている。そのため、パブリックデータでは、中間形式内部におよそ送られてくるであろうとおもわれる標準的な撮影コンディションのデータを保持し、悪意のあるユーザから著作物の権利を守る機能を付加・強化する必要がある。

また、昨今は、ディープラーニングなどの機械学習等を使い、複雑なレイアウトにも対応する研究があり、それらの研究を取り入れていくことや、国会図書館主導で開発された OCR エンジンの採用など、最新技術の導入も検討したい。

用語

ディスレクシア：文字情報を脳で処理することが正常にできず、聴覚を補助情報として必要とする高次脳障害の一種。知的レベルは問題はないが、縦書きの文字を読めず、横書きにすると読めたり、システムが音読することで読書ができるといった症例が報告されている。

AI：ここでは、AI を広い意味で扱い、画像認識技術などを含んだ自動化も含めている。ディープラーニングなどの技術を使った AI の技術に特化した場合は、機械学習という用語を用いた。

書誌情報：本のタイトル、発行年、などの付随データのこと。国会図書館のものが有名であるが、新刊の書誌データは発刊後すぐに得られないため、有償でかつ Amazon や Google が用いているものを発売元企業の協力を得て、今回利用した。