

人工知能技術を活用した動画や画像などのメディアを対象としたファイル検索システム Search Mission の事例紹介

Case study of a file search system “Search Mission” of a target of media of a movie and image, etc., using AI Technology

兼安 勉^{1*}

Tsutomu Kaneyasu¹

¹ アルファテクノロジー株式会社

¹ Alpha Technology Co., Ltd.

Abstract: As the usual provision of products and services with AI technology of explosive leap, in order to solve the problem of time and cost, and the problem of engineer shortage, many small and medium enterprises are using API services of other companies' AI technology.

In addition, based on the shift to a media medium due to normalization of smartphones, etc., and based on "work style reforms" required for today's companies, this paper will introduce system examples to be implemented to efficiently search files using AI technology.

1 はじめに

爆発的に飛躍中の人工知能技術搭載の製品やサービス提供が当たり前になる中で、時間とコスト、更に技術者不足の課題を解決するべく、多くの中小企業は他社の人工知能技術の API サービス活用を選択する。

また、スマートフォンなどの常態化によるメディア媒体へのシフト、昨今の企業に求められる「働き方改革」を踏まえ、本書では、人工知能技術を活用し、従来対象外であったファイル検索を効率的に実施するシステム事例を紹介する。

2 課題

Web 上から欲しい情報を検索する仕組みは、Google などを筆頭に浸透している。

ただし、企業内の検索には多くの課題が存在し、システム単位での検索利用となっているのが現状である。

*連絡先：アルファテクノロジー株式会社

事業統括部 製品開発課

〒150-0043 渋谷区道玄坂 2-10-7 新大宗ビル 1 号 9F

E-mail: kaneyasu.tsutomu@alphatc.co.jp

課題としては、以下 4 点となる。

- ① 資料が見つからない
- ② 色々なシステム利用の不便さ
- ③ 色々なメディアの活用拡大
- ④ 社内、社外でも検索、閲覧したい

2.1 資料が見つからない

図 1 に示すように、人は時間経過するごとに記憶したことを忘れてしまう。^{*1} 図 2 の赤枠には、企業の業務において、必要な文書を容易に取り出せない問題から生じる対処に費やす時間の割合を示す。^{*2}

1 年に換算すると、約 3 ヶ月分となる。

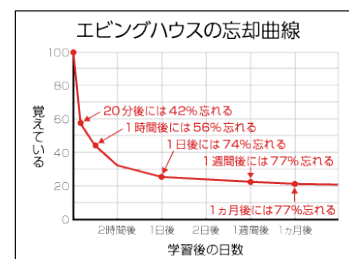


図 1：エビングハウスの忘却曲線

インフォメーションワーカーが問題への対処に費やす時間：個人の生産性

	1週間に費やす時間	費やす時間の割合	1週間に無駄にする時間	無駄にする時間の割合	企業の生産性が失われる割合
複数のファイルやフォーマットに存在する情報を1つの文書に集約する	4.1	8.2%	1.0	2.0%	1.7%
紙の文書で発生する問題や手間のかかるタスクに対処する	3.9	7.8%	1.0	2.0%	1.6%
文書を探すが見つからない	2.7	5.5%	2.7	5.5%	4.4%
文書の最新バージョンまたは正しいバージョンが見つからないか失われてしまったため、再作成する	2.3	4.6%	2.3	4.6%	3.7%
合計	12.9	26.1%	7.0	14.1%	11.4%

n = 140 (国内のインフォメーションワーカー)
Notes:
・調査方法は補遺/関連資料を参照。割合は、Figure 1の作業に費やす1週間当たりの52.4時間という回答と、従業員の81%がインフォメーションワーカーであるというデータに基づく。
・列の合計値が合わない場合は、四捨五入による誤差である。

図2：問題への対処に費やす時間

2.2 色々なシステム利用の不便さ

Amazon EC2 サービス開始、Microsoft Azure に登場により、クラウド市場が2006～2008年にかけて大幅に拡大してきた。^{*3}

以前は企業内でサーバ構築、運用するオンプレミスの形態が主流であった。

上記の流れから、当初はオンプレミスから全てクラウドへ移行し、統合システムを期待されていた。実際には、図3に示すように、システム統合が困難で、業務に適したシステム活用を目指し、オンプレミスとクラウドを混在したハイブリッドクラウドが2015年以降から徐々に割合を増やしている。^{*4}

そのため、企業内では様々なシステムの混在化は継続し、それに伴う情報連携への課題が残ったままである。

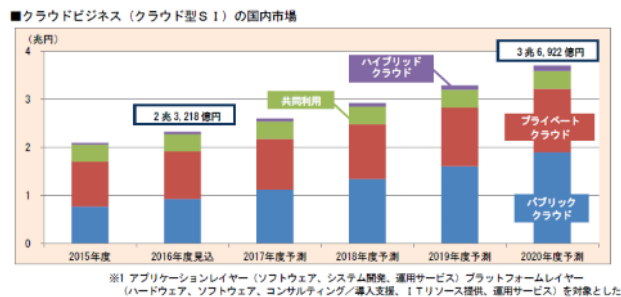


図3：クラウドビジネスの国内市場

ここで文書について示す。文書とはe文書法^{*5,*6}で定めている。

代表的な文書としては、

- ・会計帳簿
- ・契約書、見積書等
- ・振替伝票
- ・営業報告書
- ・財産目録
- ・規約等
- ・議事録等
- ・設計書

などが挙げられる。

また、e文書法における基本要件としては、

- ①見読性
- ②完全性
- ③機密性
- ④検索性

が挙げられる。

項目2.1, 2.2で重要な要件としては、上記の③機密性、④検索性が該当する。

2.3 様々なメディアの活用拡大

企業内の業務は、紙媒体からPDFやOffice文書などのドキュメントファイルの電子媒体が中心になった。昨今では、SNSの台頭、主流により、図4に示すように、動画のメディア活用が拡大している。^{*7}

今後は更に様々なメディア活用は拡大していき、ドキュメントファイルの利用率と同等になってくるとも考えられる。

ここで企業における動画活用の主なシーンは、

- ・販促、営業
- ・マニュアル
- ・カスタマーサポート
- ・社員教育
- ・採用活動

となる。

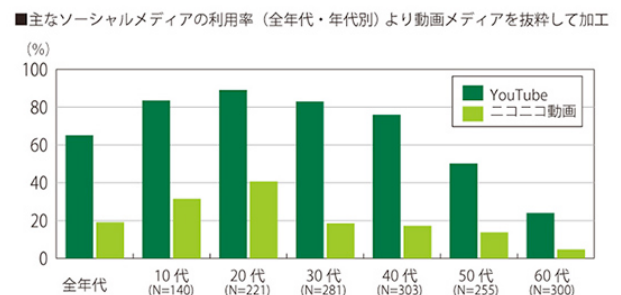


図4：SNSによる動画メディア加工

2.4 社内、社外でも検索、閲覧したい

スマートフォンやタブレットの利用率は全世代でも7割を超え、企業内でも必需品となった。

主な利用は、図5に示すように、

- ・営業活動の効率化
- ・業務時間の短縮
- ・ファイル・データの共有

などである。^{*8}

どの項目も社内、社外の場所を問わず、どこでも簡単に欲しい情報を検索し、閲覧できるといった要素を含んでいる。

また、社外からの閲覧はセキュアな環境が求められているため、項目 2.2 にも記載している基本要件の重要な点の機密性、検索性にも密接に関連する。

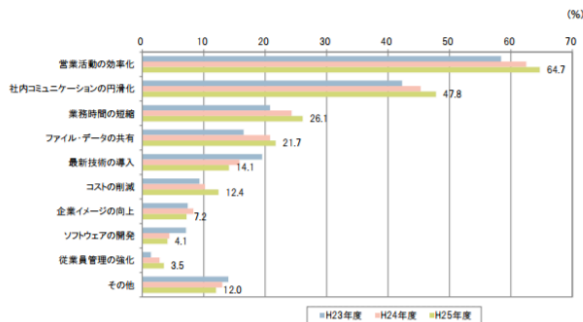


図 5: スマホやタブレットの業務利用シーン

3 提案方法

3.1 従来の検索方法

テキストを抽出できる PDF や Office 文書などのドキュメントファイルから全文テキストを抽出し、インデックス検索するために再構築したデータベースを利用する。

インデックス検索とは、以下 3 点のポイントがある。

- ①事前に文書から単語を抽出し、各単語の頻度を数値化
- ②検索キーワードとなるクエリ頻度と、上記①で抽出した文書頻度をコサイン類似度で測定
- ③コサイン類似度から定めたルールに応じて（例：高い値）、ヒット文書を抽出

インデックス検索の例としては、下図 6 が存在するとき、各文書のコサイン類似度は以下となる。

ここでは文書 3 を抽出する。

	私	あなた	年齢	性別	扉	空	雲	雨	晴
文書1	0	0	0	0	0	0	0	0.71	0
文書2	0	0	0.71	0.48	0.18	0	0	0.48	0
文書3	0.64	0.44	0	0	0.16	0.44	0	0	0
クエリ	0	0	0	0	0.35	0.94	0	0	0

図 6: インデックス検索の数値

文書 1 とクエリのコサイン類似度 =

$$0 \times 0 + \dots + 0 \times 0.35 + 0 \times 0.94 + \dots = 0$$

文書 2 とクエリのコサイン類似度 =

$$0 \times 0 + \dots + 0.18 \times 0.35 + 0 \times 0.94 + \dots = 0.06$$

文書 3 とクエリのコサイン類似度 =

$$0 \times 0.64 + \dots + 0.18 \times 0.35 + 0.44 \times 0.94 + \dots = 0.47$$

3.2 畳み込みニューラルネットワーク手

法によるタグ抽出

画像を複数のカテゴリに分類するように学習し、タグ付けを実現するには、一般的な手法である畳み込みニューラルネットワーク (Convolutional Neural Network) を利用する。

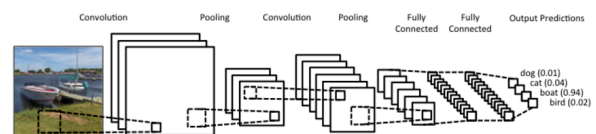


図 7: 畳み込みニューラルネットワーク

図 7 に示すように、Clarifai^{*9} が提供している畳み込みニューラルネットワークは、①畳み込み、②プーリング、③畳み込み、④プーリング、⑤全結合、⑥全結合で構成されている。

畳み込みは図 8 に示すように、画像内のピース単位のピクセル値に、フィルタとなるピクセル値を掛け、特徴のピクセルの総数で割ることで求める。^{*10}

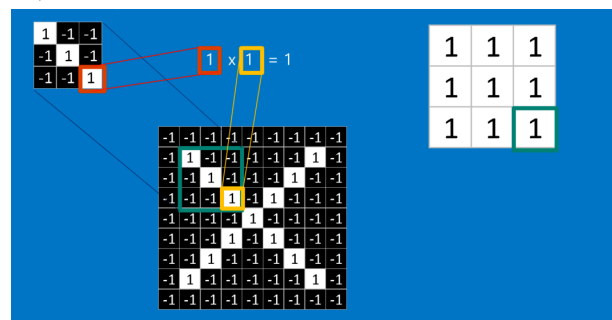


図 8: 畳み込み

プーリングは図 9 に示すように、大きな画像を重要な情報を残しつつ縮小する方法である。画像内を小さなウィンドウに区切り、区切った各々のウィンドウ内の値から最大値を求めるものである。

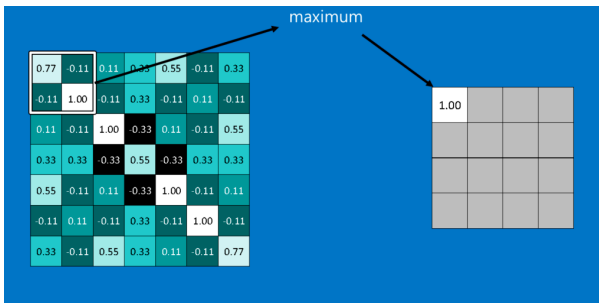


図 9：プーリング

本稿では 1024 次元の全結合層となる。

前記畳み込みニューラルネットワーク手法で学習したモデルを活用することで、図 10 に示すように、画像を入力すると、「幸せ」や「純粋」タグといった表情などを認識すること、「猫」タグといった人間以外や職業なども認識することができる。



図 10：タグ抽出

3.3 本提案の検索方法

本提案におけるインデックス検索用のデータベースへ抽出したタグを登録するフローを図 11 に示す。

アップロードしたファイルに対し、処理を行いやすいようにファイルを変換する。ここでは、ドキュメント類は PDF へ、動画類は mp4 へ、画像類は jpg とする。

次に、ファイルのフォーマットタイプ毎にサムネイル画像の抽出方法が分岐する。

PDF の場合は、先頭ページの全体をサムネイル画像で抽出する。

mp4 の場合は、動画内の各画像から抽出した解像感の高いものをサムネイル画像で抽出する。解像感での抽出方法は別途記す。

jpg、その他の場合は、画像自体、または、所定画像をサムネイル画像で抽出する。

本提案では抽出したサムネイル画像に対し、前項 3.2 で記した畳み込みニューラルネットワークモデルを用いて、認識値の高い上位 10 個のタグを抽出する。

ここでの畳み込みニューラルネットワークの事前学習で活用している情報は、人物、モノ、感情などを含めてタグ数 11,000 個となる。

PDF 以外は、タグのみでインデックス検索用のデータベースへ登録する。PDF は全文テキストを抽出し、タグと連結した情報でインデックス検索用のデータベースへ登録する。

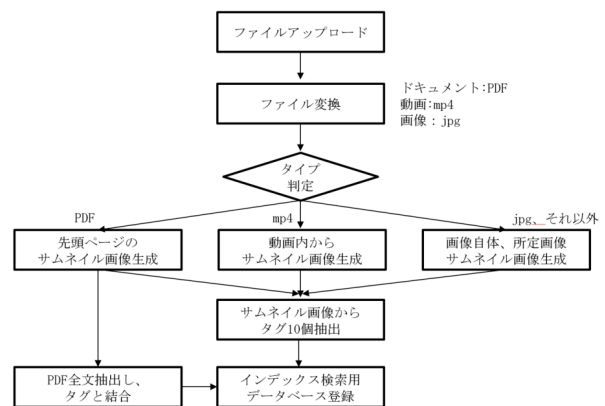


図 11：インデックス検索 DB へのタグ登録手法

解像感での抽出は下記①、②を処理する。

①画像から輝度値の 2 次元配列を生成する。

$$\text{輝度値} = 0.298912 * r + 0.586611 * g + 0.114478 * b$$

P11	P12	P13	P14	P15	P16	...
P21	P22	P23	P24	P25	P26	
P31	P32					
P41	P42					
P51						
P61						
.						
.						
.						

②横同士、縦同士の全組合せの輝度値の差の比を足し合わせて平均を取得する。

$$R[11-12] = |P11 - P12| \div (P11 + P12)$$

$$\text{SUM[横方向]} = R[11-12] + R[12-13] + \dots + R[21-22] + R[22-23] + \dots$$

$SUM[縦方向] = R[11-21] + R[21-31] + \dots + R[12-22] + R[22-32] + \dots$

評価値 = (SUM[横方向] + SUM[縦方向]) ÷ 個数

4 事例

4.1 アニメ業界

アニメ業界では背景やモノなどの素材画像を再利用して活用するケースが多い。ただし、素材画像の管理が外付けハードディスクを利用し、乱雑な状態である。また、素材画像に対しても正しいタグ付けが行われていないという課題がある。

このような課題を解決するため、本ターゲットに絞り込んだ畳み込みニューラルネットワークでカスタマイズ学習したモデルも利用することにする。

まず本提案では、図 12 に示すように、AI 技術を活用した検索システム『Search Mission (サーチミッション)』を基盤とする。Search Mission 自体は Windows Server の Web システムである。

システム構成としては、

- Web サーバ： Nginx 1.11.4
- PHP： 7.0.10
- PHP フレームワーク： CakePHP3.0
- DB サーバ： PostgreSQL 9.5.7-1

となる。

エンドユーザーは、ブラウザ (Internet Explorer 11 以上, Firefox 24.0 以上, Google Chrome 29.0 以上, Safari iOS のみ 5.0 以上) さえあれば、デバイスを問わずに操作することができる。



図 12：Search Mission 構成図

図 13 に示すように、一般モデルを活用して抽出した 10 タグ (コンセプト) だけでなく、カスタマイズしたモデルからピンポイントとなる 1 タグ (コンセプト) を合わせて抽出し、インデックス検索用のデータベースへ登録する構成する。

今回は図 14 のように 8 タグ (コンセプト) に対し、Positive, Negative 画像を各々学習した。1 タグ当たり画像 30 枚程度が推奨という中で、Positive に関して

は、十分な枚数を確保することが出来た。

結果として、本事例での認識値は低い、認識結果の中での上位タグを抽出すると、望むタグを識別することは出来ている。

ただし、今後の改善点は、

- タグ (コンセプト) の数を増やしたときの識別
 - 複数キャラクタなどが含まれているときの識別
 - カスタマイズモデルのリアルタイム更新
- が挙げられる。

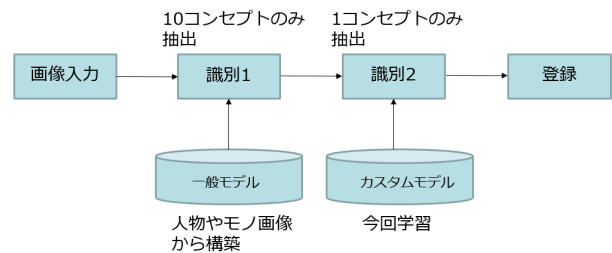


図 13：識別 (タグ抽出) 構成

コンセプト	Positive (ファイル数)	Negative (ファイル数)
キャラクタA	79	21
キャラクタAの口	14	11
刀	52	0
女性A	53	50
女性B	49	37
男性C	13	10
男性D	12	22
黒い車	28	0

図 14：学習データ数

4.2 納品 & FAQ 提供

企業には、ファイル管理システムや転送媒体の暗号化システムを利用するケースが多い。ただし、ファイル管理を社外との共有で利用することができないことや、媒体転送はあくまでも社外へ一方的に送信するだけで社外からの負担や転送方法などが存在しないといった課題がある。

また、製品紹介やFAQは別システムや別部署の担当者で管理されるため、構築のハードルが高く、大企業以外で提供されるケースは少ない。

本事例では図 15 に示すように、Explorer 風のフォルダ管理を実装している。各フォルダを社外ユーザー単位で管理し、アクセス権限を付与することで、各フォルダ内に納品物、プロジェクトに必要なデータを互いにアップすることが出来るため、社内外で

の共有を実現することが可能となる。

また、製品、会社紹介の動画を自動タグ付けでアップロードすることができるため、ユーザーが自然と他の情報を目にするような環境を構築できる。

最後に、製品に関する FAQ データは、企業内で蓄積している。Search Mission へ質問、回答の 1 グループ単位をテキストデータでアップすることで、容易に FAQ システムを構築することができる。

多くの FAQ データを人手でアップするには工数を費やしてしまうが、NTT データ製 WinActor^{*11} といった RPA (Robotic Process Automation) ツールを活用すると、数分でデータアップを完了することができる。

RPA とは、認知技術 (ルールエンジン、機械学習、人工知能等) を活用した、主にホワイトカラー業務の効率化、自動化の取組みである。人間の補完として業務を遂行できることから、仮想知的労働者 (Digital Labor) とも言われている。^{*12}



図 15 : Search Mission UI

むすび

本稿では、企業内の検索に存在する 4 個の課題、①資料が見つからない、②色々なシステム利用の不便さ、③様々なメディアの活用拡大、④社内、社外でも検索、閲覧したい、の各々について整理した。

前記課題に対して、人工知能技術で一般的に利用される畳み込みニューラルネットワーク手法、インデックス検索手法の基礎について説明した。その後、ドキュメントだけでなく、画像や動画といったメディアも含めた検索方法におけるタグ登録フローを中心に紹介した。

最後に、アニメ業界、納品&FAQ 提供の 2 事例を活用して、人工知能技術搭載のファイル検索システムを紹介した。

今後としては、

- ・タグ (コンセプト) の数を増やしたときの識別
 - ・複数キャラクタなどが含まれているときの識別
 - ・カスタマイズモデルのリアルタイム更新
- の改善点に加え、動画全体に渡るタグ抽出や、音声に対するタグ抽出なども実現していきたい。

謝辞

本稿作成にあたりましては、某社の情報の御提供、ならびに種々の御助言を賜りました。ここに深甚なる謝意を表します。

参考文献

- [1] フリージュニアアカデミー, <http://free-academy.jp/junior/index.php?エビングハウスの忘却曲線>
- [2] 日本国内のインフォメーションワーカーが抱える生産性ギャップを埋める: IT 部門の新たな課題と機会, IDC, 6 月, 2012
- [3] 丸山不二夫, クラウド時代の始まり, <https://thinkit.co.jp/article/1024/1>, 2009
- [4] 2017 クラウドコンピューティングの現状と将来展望, 株式会社富士通キメラ総研, 2017
- [5] e-文書法によって電磁的記録が可能となった規程, 経済産業省,
- [6] 文書の電磁的保存等に関する検討会, 文書の電子化の促進に向けて, 経済産業省, pp.48 - 64
- [7] 平成 26 年情報通信メディアの利用時間と情報行動に関する調査 報告書, 総務省, pp84, 2014
- [8] 平成 26 年度我が国情報経済社会における基盤整備報告書, 経済産業省, 2014
- [9] Clarifai, <https://www.clarifai.com/>
- [10] 「畳み込み」と「プーリング」, <http://postd.cc/how-do-convolutional-neural-networks-work/>
- [11] WinActor, <https://winactor.jp/>
- [12] Wikipedia, <https://ja.wikipedia.org/wiki/ロボティック・プロセス・オートメーション>