

創造活動のための情報資源の能動的獲得と集積機能の検討

河間勇氣^{†1, †2} 福谷遼太^{†1, †2} 高橋秀幸^{†1, †2} 木下哲男^{†1, †2}

概要: 携帯端末の高性能化・高機能化、クラウド技術の発展によって、いつでも、どこでも創造活動が可能になりつつある。創造活動においては、PC、携帯端末などを利用して作成した資料、Web から収集した情報、メールのやり取り、音声、画像、テキストのメモなどの膨大かつ多様な情報資源を参考にしながら、アイデアの発想や成果物の制作を行うことが多い。しかし、過去の創造活動時に利用した情報資源の保存先や参照先、ある作業時に利用した情報資源を体系的に人が記憶することは困難であり、創造活動時に利用する端末やソフトウェアも多様化しているため、特定の情報資源を探し出すことに時間を要してしまうことがある。我々は、この問題を解決するために、過去の創造活動の記憶想起を支援する、能動的情報資源に基づく構造体獲得支援機構を提案する。この機構では、情報資源同士が、創造活動時のソフトウェア、情報資源の利用状況、メタデータなどに基づいて能動的に関連し合い、創造活動の構造体を形成する。本稿では、提案機能とその試作について述べる。

キーワード: 創造活動、分散処理、マルチメディア、想起支援、エージェント

Active Acquisition and Accumulation Function of Information Resources for Creative Activity

YUKI KAWAMA^{†1, †2} RYOTA FUKUTANI^{†1, †2}
HIDEYUKI TAKAHASHI^{†1, †2} TETSUO KINOSHITA^{†1, †2}

Abstract: With the advancement of mobile device and cloud computing, today, we can conduct a creative activity anytime and anywhere. Usually, we refer enormous and diverse information resources such as created materials, web pages, e-mails, voices, images and memos to create something. However, it is difficult for people to memorize all saved destination of information resources. Furthermore, it may take time to find a specific information resource because of the diversification of devices and software. In this paper, we will discuss about information organize support function based on active information resources. This function supports information resources to construct relationships based on software, utilization state of information resources, metadata that a user used in a creative activity. Then the function makes a structure of information resources used on a specific creative activity.

Keywords: Creative activity, Distributed processing, Multimedia, Recall support, Agent

1. はじめに

研究から芸術まで様々な創造活動は、基本的に、制作物や参考資料など、過去に制作・収集されたもの(情報資源)に基づいて行われることが多い。例えば、我々が新たに、研究発表に向けて発表用スライドを制作するとき、情報資源として以前制作した文書や、参考にした学術論文を確認しながら、自身が過去に行ってきた研究の内容を想起することで、スライド制作のための着想を得る。創造活動で情報資源を活用するにあたり、近年は、情報の記録や閲覧に役立つPCや携帯端末が積極的に利用されている。さらに、携帯端末の小型化・高性能化やクラウドストレージの普及により、時間と場所を問わず、情報資源の確認や活用ができるようになった。

我々は普段、後日必要に応じて適当な情報資源を確認できるように、情報資源をフォルダに格納するなど、体系的に整理する。しかしながら、情報資源は、それぞれ情報の

所在(携帯端末や Web)や、情報の種類(メール、メモ、音声、画像)が多岐に渡る上に、膨大である。そのため、各情報資源の保存先を一つ一つ手作業で指定し、体系的に整理する作業は、困難かつ負担が大きい。

また、一般的に人間の記憶は、経験した一つの出来事に対して、その出来事が起きたときの活動状況(どこで何をしていたかなど)が紐づくと言われている。そして、現在の出来事に近い経験を過去にしている場合、その経験を想起できれば、過去の経験を活かして行動することができる。

そのため、特定の創造活動時に扱った情報資源を活動時の状況とともに整理し、利用者の現在の活動状況に応じて、図1に示す例のように提示できるシステムを開発することができれば、過去の創造活動の想起を促し、その時の経験や利用した情報資源を活かした創造活動の支援をすることができる。

上述した創造活動支援システムを実現するためには、情報資源を活動時の状況とともに整理する必要がある。しか

†1 東北大学電気通信研究所
Research Institute of Electrical Communication, Tohoku University
†2 東北大学大学院情報科学研究科

Graduate School of Information Sciences, Tohoku University

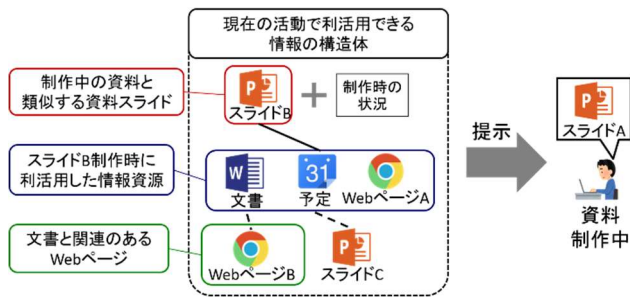


図 1 想定する創造活動支援

しながら、前述の通り、情報資源を手で適切に整理する作業は利用者に大きな負担を与える上、見落としなどのヒューマンエラーによって整理に失敗する可能性もある。このため、システムには自動的な整理を行う機能が求められるが、その機能を実現するためには、利用者の作業状況を推定し、作業状況に応じて情報資源を整理する仕組みが必要になる。さらに、整理した情報資源は、利用者の状況・要求に応じた提示のため、利用した際の状況と共に、体系的に管理されなければならない。

本稿では、これらの仕組みを有する情報資源の構造体獲得支援機構の提案と、その機能検証実験について述べる。

2. 関連研究と技術的課題

2.1 関連研究

利活用状況を基に情報資源を整理・提示する様々な研究がある。過去に閲覧した Web ページの再訪問を支援する研究[1]では、利用者が Web ページを参照していた際の時間と場所、活動状況を記録し、コンテキストツリーを作成することで、ページ内のキーワードだけでなく、「いつ」「どこで」「何をしているときに」といったコンテキストキーワードを用いた Web ページ検索を可能とするシステムを提案している。このシステムでは、ページ参照時に同時に実行していたコンピュータプログラムのフォーカス時間・頻度などから、利用者の活動状況を推定する。また、同様に Web ページの再訪問を促す研究があり、活動状況を基にした Web ページ探索を実現しているが、再訪問支援の対象範囲は、Web ページにとどまる[2][3]。

ある文書の制作など、同一の活動で利用したファイルの発見支援を目指す研究[4][5]では、ファイルの添付元ファイルとの関係やフォーカスしているアプリケーションの遷移などのログを用いることで、多様なファイル間の距離計算を行い、クラスタリングする手法を提案している。さらに、各クラスタで利用者が最も長く操作していたファイルを、利用者の作業の目的と仮定することで、作業の種類に応じた整理も可能である。しかしながら、ファイルの内容の関連性については考慮されておらず、単一の情報端末での利用に特化した手法であるため、複数の情報端末を利用した創造活動の支援を実現することが難しい。

一方、創造活動における情報資源の管理・参照に関する研究がある。例えば、研究に関して議論した内容の効率的な整理・記録から、発表資料の作成まで総合的に支援するためのシステムを提案する研究[6][7]や、アイデア等のノートに関連度に応じて配置するインターフェースを用いて、利用者の記憶の想起を支援する研究[8]がある。これらの研究は、利用者に対し、思考の整理や情報資源の想起を促す環境を提供することができるが、情報資源の整理を人手でうまく行うことが前提であるため、情報資源の整理負担を利用者に強いることとなる。

以上より、既存の情報資源を整理・提示する手法は、多種の情報資源・複数の端末を利用した活動に対応すること、利用者が整理に要する負担を軽減することの両立が実現できていない。このため、本研究が対象とする創造活動の支援に適用することは困難である。

2.2 技術的課題

利用者が創造活動を行った際に利活用した情報資源を、利用した状況や意図を反映した形で整理・構造化し、後に利用者の状況に応じた提示ができるよう管理する仕組みを実現するには、以下の技術的課題がある。

まず、創造活動においては複数の端末上で多様な形式・内容の異なる様々な情報資源を利用する。しかし、形式の異なる様々な情報資源について、内容の類似性などを考慮した関連度を一意に計算することは困難である。さらに、複数の端末での活動状況を考慮して整理するために、利用した各端末での活動状況を取得する必要がある。すなわち、異なる種類の情報資源間でも適切に関連度を計算できる仕組みと、端末の種類に依存しない活動状況の取得手法が必要となる。

また、本研究における最終的な目的は、現在の利用者の活動状況に応じて、適切な情報資源の提示を行うことである。これを実現するためには、特定の創造活動で利用した情報資源をまとめるだけではなく、「何を目的とした活動であったか」などの活動状況を効率的に管理できるとよい。このため、整理した情報資源を、利用した際の活動状況と共に、構造化として体系的に管理可能な仕組みが必要となる。

本研究では、上述した技術的課題を解決するため、情報資源を活用した際のコンテキスト情報やメタデータを基に、個々の情報資源が自律的に他の情報資源との関連性を算出し、活動内容に応じて情報資源同士を紐づけ合うことで、創造活動時の利用状況を反映した情報資源の整理・構造化が可能な情報資源の構造体獲得支援機構を提案する。

次章にて、本提案の概要と、これを実現する機能の詳細について述べる。

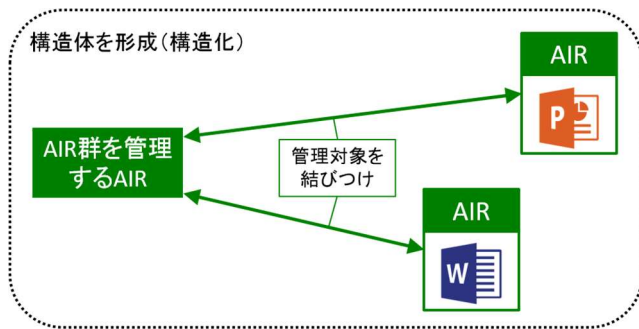
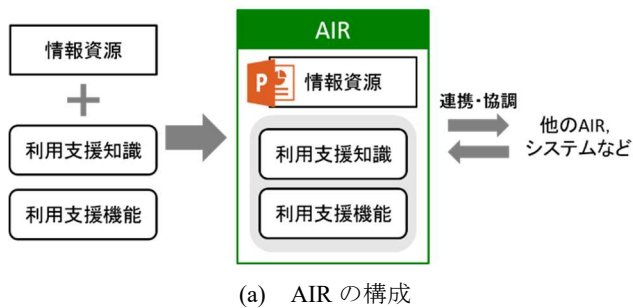


図 2 AIRの概要

3. 情報資源の構造体獲得支援機構

3.1 構造体獲得支援機構の概要

本研究では、情報資源の活動状況に応じた柔軟な関連性の算出・構築を実現するシステム技術として、各情報資源を能動的情報資源(AIR: Active Information Resource)[9]として構成するアプローチを導入する。AIRは、図2(a)に示すように情報資源を管理する機構であり、自らが管理する情報資源の内容に応じた知識を持ち、情報資源の操作や、他のAIR・システムに能動的に働きかける機能を持つ。また図2(b)にあるように、分散環境上の複数のAIRを管理するAIRを生成し、それらを構造体としてまとめること(構造化)が可能である。これにより、複数の情報資源の集合体を管理・運用する手段を得ることが可能となり、情報資源の効率的な利活用に繋がる。また、AIR化された情報資源と知識は、AIRが動作する仮想的な空間において管理される。以下、この空間をAIR管理スペースと呼ぶ。

情報資源の構造体獲得支援機構の概要を図3に示す。本機構は、以下の4つの基本機能からなる。

(1) AIR化機能

それぞれの情報資源を管理するAIRを生成し、自律的な処理が可能な機構とする機能。

(2) 活動状況獲得機能

利用者が情報資源を利活用した際の状況を取得し、知識として保持する機能。

(3) 関連性計算・構築機能

(1),(2)で得た知識を基に、情報資源間の関連度を計算し、関連性を構築する機能。

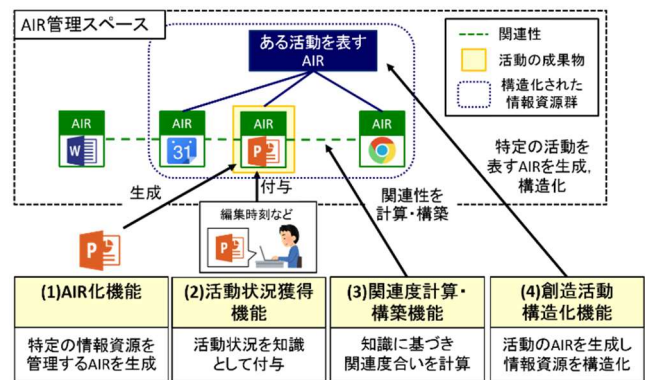


図 3 情報資源の構造体獲得支援機構の概要

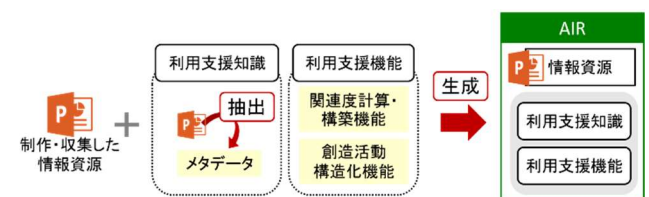


図 4 AIR化機能

(4) 創造活動構造化機能

活動の内容を推定し、活動で利用した情報資源を構造化する機能。

以下、本機構の処理の流れを述べる。まず、利用者が創造活動において創造した成果物や、参考にした資料などの情報資源からメタデータを抽出し、それを利用支援知識とすることで、情報資源を管理するAIRの生成を行う。ここで生成されるAIRは、それぞれが管理する情報資源の内容に基づき、自律的に他のAIRと協調・連携する機能を持っている。また、活動状況獲得機能により継続的に利用者の作業状況を監視することで、活動状況を取得し、これを知識としてAIRに付与する。次に、得られた知識を基に、個々のAIRがそれぞれの管理する情報資源の関連度を計算し、関連のあった項目内容とともに関連性として保持する。この関連性は、新たなAIRの生成や利活用状況の更新のたびに算出され、更新されていく。最後に、利用者が成果物を創造した際の活動状況から、活動の内容を推定し、その際に利用した情報資源を結び付けて構造体を作成することで、活動状況に応じた情報資源の構造化を行う。以下、各機能の詳細について述べる。

3.2 AIR化機能

図4にAIR化機能の概要を示す。AIR化機能は、利用者が過去の創造活動で制作した成果物や、参照した参考資料などの様々な情報資源から自動でメタデータを抽出し、利用支援知識に変換することで、情報資源を管理するAIRを生成する。ここで創造活動の成果物とは、利用者の活動における最終目標としての制作物のみならず、参考資料の内容整理のために記述したメモや、自らのスケジュールを管

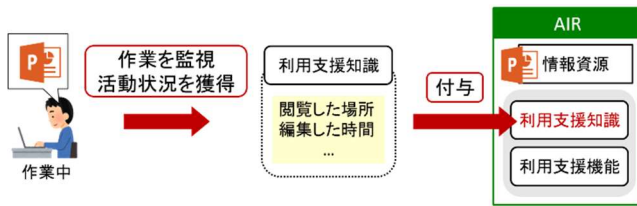


図 5 活動状況獲得機能

表 1 獲得を行う活動状況

媒体	手法	項目
特定の端末	端末専用のシステム	情報資源を閲覧・編集した期間
AIR 管理スペース	Web アプリケーション等	情報資源にアクセスした際の時間・端末・場所、前後にアクセスした情報資源

理するためカレンダーに記載した予定など、創造活動の途中で生成した様々な情報資源を含むものであり、その時々利用者の活動の目的を内包したものである。抽出するメタデータの例としては、情報資源が生成された日時や、作成者、内容の要点などが挙げられるが、情報資源の種類によってその形式は異なるため、本機能ではあらかじめ AIR が保持できる形式を定義し、メタデータをその形式へ変換することで知識とする。

さらに、生成される AIR は利用支援機能として関連度計算・構築機能と創造活動構造化機能を持つ。これらの機能の詳細についてはそれぞれ 3.4 節、3.5 節にて述べる。

3.3 活動状況獲得機能

図 5 に示すように、活動状況獲得機能では、利用者が創造活動を行っている間、どのように情報資源を利用していたかを取得・記録し、AIR の知識へと変換することで活動状況を獲得する。本稿では、複数端末の利用にも対応した高精度な活動状況の獲得を実現するため、複数の活動状況獲得手法を併用する。本稿において獲得を行う活動状況の内容を表 1 に示す。獲得を行う項目は、獲得時の媒体に基づいて、大別して「特定の端末より得られる活動状況」、「AIR 管理スペースより得られる活動状況」の 2 種類に分類される。

「特定の端末より得られる活動状況」は、利用者が情報資源を利用するために使用している端末 (PC, スマートフォン等) のプロセス情報を監視・取得することで得られる知識であり、例えば、情報資源を閲覧していた期間や、編集していた期間 (閲覧中に、情報資源の容量が変化した期間) などが該当する。これらの知識を取得するには各端末に対応した専用のプログラムを実行しておく必要があるが、利用者が何時から何時まで情報資源を利用していたかなど、作業状況を細かく入手できる。

一方、「AIR 管理スペースより得られる活動状況」は、利用者が AIR 管理スペース上の情報資源に、AIR 管理スペース

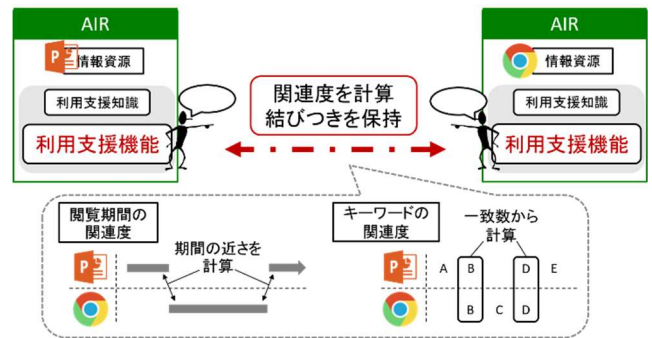


図 6 関連度計算・構築機能

ス閲覧用のユーザインタフェースよりアクセスした際の状況を取得することで得られる知識であり、情報資源にアクセスした時間・端末・場所や、前後にアクセスした情報資源の情報などが該当する。これらの知識は情報資源にアクセスした瞬間のみ取得可能であり、例えば情報資源にアクセスしたのち、いつまで利用していたかといった状況を把握することはできないが、AIR 管理スペースの閲覧に用いるユーザインタフェースを WEB アプリケーションにすることで端末の種類を問わず取得可能になる。

本稿では、前述した 2 種類の項目を獲得することで、複数端末の利用にも対応した高精度な活動状況の獲得を実現する。

これらの知識はそれぞれ、利用された各情報資源を管理する AIR の知識として保持し、後述する関連度計算や構造化に用いる。

3.4 関連度計算・構築機能

図 6 に関連度計算・構築機能について示す。本機能では、各 AIR がそれぞれ、保持する知識を基に「どの情報資源を同時に利用していたか」、「どの情報資源が内容的に類似しているか」といった関連度を計算し、算出された関連度が高い情報資源同士を結び付けることで、関連性の構築を行う。情報資源間に関連性を構築することにより、例えば、共通のテーマを扱う論文などが結びつき、後に利用者が創造活動を行う際の参考資料などの情報資源探索に要する負担を大幅に軽減することができる。

具体的な関連度計算の手法について述べる。まず、関連度計算の際に、新しく生成された AIR が、自らが持つ知識を他の AIR に対して送信する。知識を受け取った AIR は、自身の保持する知識と受け取った知識を比較し、算出式に従って関連度を算出する。そして、算出された関連度および類似性のあった項目を、知識を送信してきた AIR に返す。また、関連度は各情報資源の利活用状況によって変化するため、定期的に再計算し、関連度の値を更新する。

関連度は、情報資源を利用した期間と、内容的な類似性を考慮して算出する。また、片方の情報資源がカレンダーの予定だった場合、もう片方が予定した期間に利用した情報資源であるかどうかを考慮して算出する。具体的には、

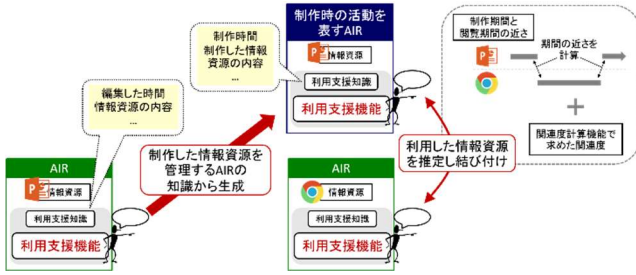


図 7 創造活動構造化機能

関連度は情報資源を利用した期間の重複度 t_{use} , 生成日時の近さ t_{create} , キーワードの一致度 k , 予定との一致度 $r_{schedule}$ により構成される. 関連度計算を行う対象の AIR を $air1, air2$ としたとき, まず t_{use} を求めるため, 観測期間の集合 $\{t_1, t_2, \dots, t_n\}$ (t は秒刻みに記録されているものとする) において, ある AIR を閲覧していた期間 $T_a(air)$ を以下の式で定義する.

$$T_a(air) = \{e_i | e_i = n_i \cdot t_i\} \quad (i = 1, 2, \dots, n)$$

$$n_i = \begin{cases} 0 & (t_i \text{ に } air \text{ を閲覧していた場合}) \\ 1 & (t_i \text{ に } air \text{ を閲覧していなかった場合}) \end{cases} \quad (1)$$

このとき閲覧期間の重複度 t'_{use} は以下の式により表される.

$$\begin{cases} n_{kl} = 0 & (t_{threshold} < |T_a(air1)_k - T_a(air2)_l|) \\ n_{kl} = 1 & (0 \leq |T_a(air1)_k - T_a(air2)_l| \leq t_{threshold}) \end{cases}$$

$$t'_{use}(air1, air2) = \frac{\sum_{k=1}^i \sum_{l=1}^j n_{kl}}{|T_a(air1)| \cdot |T_a(air2)|} \quad (2)$$

$(i = 1, 2, \dots, |T_a(air1)|)$
 $(j = 1, 2, \dots, |T_a(air2)|)$

ここで $t_{threshold}$ は閾値となる時間の長さを指す.

次いで各 AIR にアクセスした日時の一致度 t''_{use} を, 以下の式により定義する.

$$\begin{cases} n_{kl} = 0 & (t'_{threshold} < |T_o(air1)_k - T_o(air2)_l|) \\ n_{kl} = 1 & (0 \leq |T_o(air1)_k - T_o(air2)_l| \leq t'_{threshold}) \end{cases}$$

$$t''_{use}(air1, air2) = \frac{\sum_{k=1}^i \sum_{l=1}^j n_{kl}}{|T_o(air1)| \cdot |T_o(air2)|} \quad (3)$$

$(i = 1, 2, \dots, |T_o(air1)|)$
 $(j = 1, 2, \dots, |T_o(air2)|)$

ここで $t'_{threshold}$ は閾値となる時間の長さを, $T_o(air)_k$ は AIR にアクセスした日時の集合を指す.

このとき, t_{use} は係数 α, β を用いて以下の式により表される.

$$t_{use}(air1, air2) = \alpha t'_{use}(air1, air2) + \beta t''_{use}(air1, air2) \quad (4)$$

続いて t_{create} は AIR の生成された日時 $t_c(AIR)$ を用いて以下の式により表される.

$$\Delta t_c(air1, air2) = |t_c(air1) - t_c(air2)|$$

$$t_{create}(air1, air2) = \begin{cases} \frac{1}{\Delta t_c(air1, air2)} & (\Delta t_c(air1, air2) < 1) \\ 1 & (\Delta t_c(air1, air2) > 1) \end{cases} \quad (5)$$

次に k は, $air1, air2$ の保有するキーワードの数 p, q と一

表 2 活動を表す AIR の持つ知識の例

項目	内容
制作した情報資源	ゼミ発表資料
情報資源のキーワード	情報資源, 想起支援, 創造活動
制作期間	2017-09-29T07:51:00Z 2017-09-29T07:52:00Z ...

致したキーワードの数 k_{match} を用いて以下の式により表される.

$$k = \frac{k_{match}}{\min(p, q)} \quad (6)$$

また, $air1, air2$ のいずれかがカレンダーに記載した予定である場合のみ, $r_{schedule}$ の計算を行う. ここで, $air1$ がカレンダーの予定であるとき, $air2$ の閲覧期間のいずれかが, 予定された期間 $[\Delta t_1, \Delta t_2, \dots, \Delta t_i]$ に含まれる場合のみ $r_{schedule} = 1$ とする. それ以外の場合には $r_{schedule} = 0$ とする.

最後に, 関連度 R の計算式を示す. なお, $\gamma, \delta, \epsilon, \epsilon$ は各要素の重み付けのための係数である.

$$R = \gamma t_{use} + \delta t_{create} + \epsilon k + \epsilon r_{schedule} \quad (7)$$

3.5 創造活動構造化機能

創造活動構造化機能は, 関連性構築とは異なり, 利用者が行った特定の活動の内容を推定した上で, 活動のために利用した情報資源を構造化してまとめることを目的とした機能である. 本稿における創造活動構造化とは, 例えば, 利用者が文書 A の制作を行ったとき, 制作した期間や成果物の情報から「文書 A を制作した活動」を定義し, 文書 A の制作時に利用した Web ページなどを「文書 A を制作した活動」に結びつけて, まとめて管理することを指す.

具体的な構造化の流れを図 7 に示す. 構造化にあたり, はじめに, 利用者が過去に制作した成果物の特定を行う. 本研究では, 取得した活動状況から利用者が編集していた情報資源を特定し, 編集した期間における成果物とする. 次に, その成果物を制作した際の期間や場所などの活動状況から「成果物制作時の創造活動を表す AIR」を生成する. ここで生成される AIR が持つ知識は, 表 2 に示す通りであり, 主に成果物を管理する AIR が持つ知識を譲与する形となる.

最後に, 生成した「活動を表す AIR」と各情報資源を管理する AIR の関連度を計算し, 一定値以上の関連性があったものを結び付けることにより, 活動時に利活用していた情報資源の構造化を行う. このとき導出される関連度 R_5 は, 成果物を編集していた期間と情報資源を利用した期間の重複度である t_{edit} と, 式(6)の関連度 R により構成される. 「活動を表す AIR」を air_{work} とし, 構造化される情報資源を管理する AIR を $air3$ としたとき, 観測期間の集合 $\{t_1, t_2, \dots, t_n\}$ における成果物を編集していた期間 $T_e(air)$ を以下のように定義する.

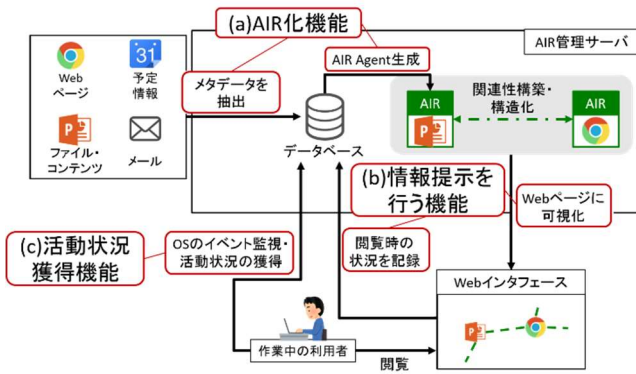


図 8 試作システムの概要

$$T_e(air) = \{e_i | e_i = n_i \cdot t_i\} \quad (i = 1, 2, \dots, n)$$

$$n_i = \begin{cases} 0 & (t_i \text{ に } air \text{ を編集していた場合}) \\ 1 & (t_i \text{ に } air \text{ を編集していなかった場合}) \end{cases} \quad (8)$$

このとき成果物編集期間と AIR3 閲覧期間の一致度 t'_{edit} は以下の式により表される。

$$\begin{cases} n_{kl} = 0 & (t_{threshold} < |T_e(air_{work})_k - T_a(air3)_l|) \\ n_{kl} = 1 & (0 \leq |T_e(air_{work})_k - T_a(air3)_l| \leq t_{threshold}) \end{cases}$$

$$t'_{edit}(air1, air2) = \frac{\sum_{k=1}^i \sum_{l=1}^j n_{kl}}{|T_e(air_{work})| \cdot |T_a(air3)|} \quad (9)$$

$$(i = 1, 2, \dots, |T_e(air1)|)$$

$$(j = 1, 2, \dots, |T_a(air2)|)$$

次いで成果物編集期間と $air3$ にアクセスした日時の一緻度 t''_{edit} を、以下の式により定義する。

$$\begin{cases} n_{kl} = 0 & (t'_{threshold} < |T_e(air_{work})_k - T_o(air3)_l|) \\ n_{kl} = 1 & (0 \leq |T_e(air_{work})_k - T_o(air3)_l| \leq t'_{threshold}) \end{cases}$$

$$t''_{edit}(air_{work}, air3) = \frac{\sum_{k=1}^i \sum_{l=1}^j n_{kl}}{|T_e(air_{work})| \cdot |T_o(air3)|} \quad (10)$$

$$(i = 1, 2, \dots, |T_e(air_{work})|)$$

$$(j = 1, 2, \dots, |T_o(air3)|)$$

このとき t_{edit} は以下の式により表される。

$$t_{edit}(air_{work}, air3) = \alpha t'_{edit}(air_{work}, air3) + \beta t''_{edit}(air_{work}, air3) \quad (11)$$

最後に、構造化に用いる関連度 R_s は以下の式により表される。なお、 θ, μ は係数である。

$$R_s = \theta t_{edit} + \mu R \quad (12)$$

生成された「活動を表す AIR」と構造化された情報資源は、利用者の過去の活動状況を示す断片でもあるため、利用者に過去の活動内容の想起を促すことが可能となる。

4. 試作システムの設計と実装

4.1 試作システムの設計

提案に基づき実装した試作システムの概要を図 8 に示す。本試作システムは、サーバ上で稼働する 2 つの機能と、利用者が創造活動において最も頻繁に活用する端末（本システムでは PC を想定）上で稼働する活動状況獲得機能の 2

種により構成される。

サーバ上で稼働する機能には、AIR 化機能(図 8(a))と、利用者へ情報提示を行う機能(図 8(b))が含まれる。また AIR は、サーバ上で動作するマルチエージェントシステムとして実装を行い、AIR を管理するスペースをエージェントワークスペースとして実装する。一方、利用者が活用する端末上で稼働する活動状況獲得機能(図 8(c))は、端末上で実行されている OS のイベント処理を取得・処理してサーバへと送信する。

本システムにおける処理の流れを簡単に述べる。まず、利用者がクラウドストレージなどに保管している情報資源から、サーバ上の AIR 化機能がメタデータの抽出・加工を行い、知識としてデータベースに保存し、その知識を基に AIR の生成を行う。この時点で AIR は、他の生成された AIR と相互に知識の送受信を行い、関連度計算・構築および構造化を実行する。この処理は、AIR に与えられた知識が変化する度に実行される。利用者は、サーバに構築された可視化インタフェースにより、いつでも AIR およびその関連度や構造体を閲覧することが可能である。また、利用者が可視化インタフェースで情報資源にアクセスしたとき、その日時や場所・利用端末といった利用状況の情報がサーバに送信され、AIR の持つ知識が更新される。さらに、利用者が PC 上で AIR 化された情報資源を利用して作業を行う際も、その作業状況が活動状況獲得機能によって取得され、AIR の持つ知識が更新される。

4.2 試作システムの実装

試作システムの実装について述べる。本システムは、AIR 化機能、情報を提示する機能、AIR を実現するマルチエージェントシステム、利用者が活用する端末上の活動状況獲得機能により構成される。

AIR 化機能は、提案における AIR 化機能に該当するものであり、今回は主に Java で実装した。本機能では、情報資源のメタデータ抽出のため、定期的に各種 API を用いて外部サービス・アプリケーションにアクセスする。今回は、Google Drive API[10]、Google Calendar API[11]、Gmail API[12] を利用し、対象のサービスからメタデータ抽出を行う。Google Drive に保存されたファイルは、ドキュメント分析・メタデータ抽出ツールキット Apache Tika[13]を用いて解析し、メタデータを抽出する機能を実装した。このとき、ファイルの内容からキーワードを抽出するため Java 製の形態素解析ライブラリ lucene-gosen[14]を利用し、特徴的な頻出語を取得する。また、抽出したデータは、AIR が保持・利用できる知識へと変換を行い、データベース(MongoDB[15])に格納する。最後に、得られた知識を基に、Java で実装したマルチエージェントシステム上に AIR エージェントを生成するようメッセージを送信する。また、Web ページを AIR 化する Google Chrome の拡張機能(Chrome extension[16])を実装した。実装した Web ページ AIR 化機能の動作例を

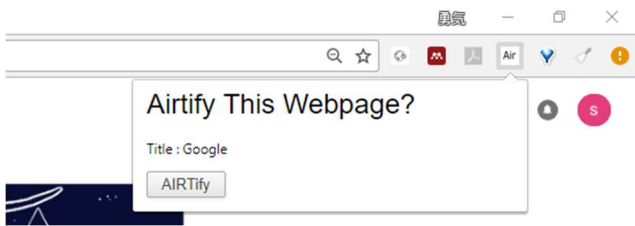


図 9 Web ページ AIR 化機能の動作例

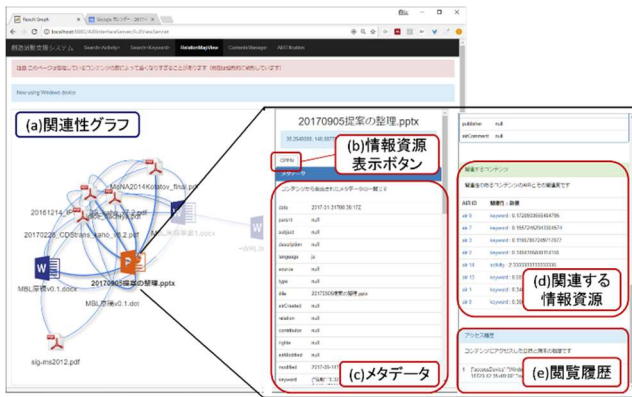
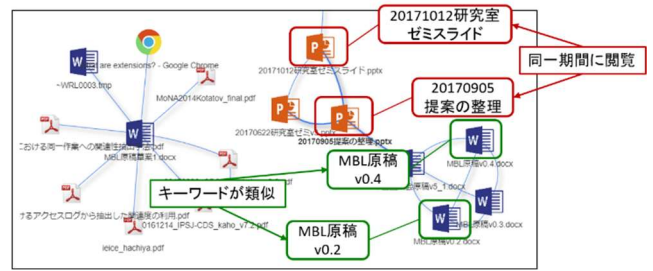


図 10 情報提示を行う機能による提示例

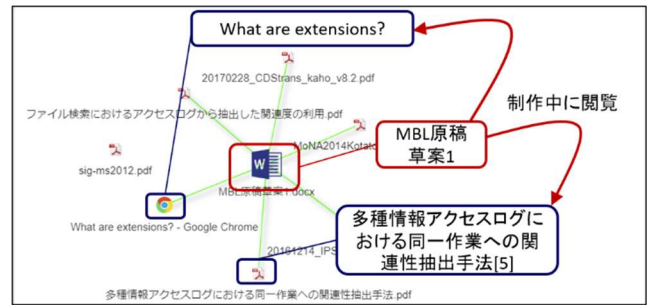
図 9 に示す。この機能は、使用中の Google Chrome に付与することで利用可能となる。具体的には、AIR 化したい Web ページを開いている状態で本機能呼び出し、「AIRtify」ボタンを押すことで Web ページの各データがサーバに送信される。その後の知識への変換、データベースへの格納が同様に実行される。

情報提示を行う機能としては、関連度計算・構築および構造化の結果を可視化し、利用者に提示すると同時に、提案における活動状況獲得機能の「AIR 管理スペースの利用状況より得られる活動状況」を取得する。本機能は、Java Servlet, JSP により実装した。本機能による提示画面の例を図 10 に示す。図 10(a)では、関連性構築・構造化された AIR の一覧が提示される。なお、グラフ内のアイコンノードは各 AIR を、青線のエッジは構築された関連性を示す。図 10(b)のボタンからは選択した情報資源にアクセスすることが可能であり、ここでアクセスした際の時刻、場所、端末がサーバに送信される。サーバに送信された情報は、それぞれの情報資源を管理する AIR に送られ、知識として付与される。また、図 10 中の(c), (d), (e)は選択した情報資源の内容や、利用日時などの情報を提示している。

利用者が活用する端末上の活動状況獲得機能は「特定の端末より得られる活動状況」を取得するため、Windows OS を対象として、C#により実装した。OS 上のイベントの取得には Windows API[17]を用い、一定期間おきにアクティブ状態になっているウィンドウの情報、ウィンドウで開いている情報資源（ファイル）の名称および情報資源の容量の変化を監視する。これにより得られたデータを基に、現在閲覧している情報資源、編集している情報資源を判断し、



(a) 情報資源の関連性グラフの例



(b) 情報資源の構造化グラフの例

図 11 提示された関連性・構造化グラフの例

サーバに送信する。送信された情報はそれぞれの情報資源を管理する AIR に送られ、知識として AIR に付与される。

5. 実験と評価

5.1 実験概要

提案機能の評価するため、一名の大学院生を利用者とし、試作したシステムを用いて実験を行った。本実験では、利用者が 2017 年 10 月 16 日から 2017 年 10 月 21 日までの期間に創造活動で扱った情報資源 54 個を、本機能を用いて利用した活動ごとに整理・構造化する。なお、扱った情報資源の内訳は、利用者が過去に収集・制作した資料が 49 個、期間中に制作を行った資料が 5 個であり、具体的には、参考にした論文、研究室のゼミ発表のため制作したスライドや、本原稿の草案などである。利用者は普段の創造活動（研究活動）において図 8(c)の機能を持つアプリケーションを搭載した Windows PC を使用し、外出先ではスマートフォンで情報資源の閲覧などを行った。また、情報資源にアクセスする際には、利用者は常に AIR 管理スペースを可視化した図 10 の情報提示インターフェースを用いることとした。

実験の結果、関連性構築が行われたグラフの例を図 11(a) に示す。図中のアイコン型のノードは取り込まれた情報資源を、エッジは構築された関連性を示す。グラフの内容の例としては、利用者が「20170905 提案の整理」と「20171012 研究室ゼミスライド」を同じ期間に閲覧し、同一のキーワードとして「研究」、「ゼミ」等を保有しているため、両情報資源の間に関連性を示すエッジが構築されたことが確認できた。また、構造化された情報資源のグラフの例を図 11(b)に示す。図中に示されているように、特定の資料を制作した活動は大きめのアイコンによって表示される。また、

表 3 各活動で利用した情報資源の正解数

	活動の種類	扱った情報資源の数
A	スライドの制作	3
B	文書の執筆	8
C	文書の執筆	14
D	文書の執筆	13
E	図の制作	7

表 4 実験結果

適合率	再現率	F 値
0.833	0.436	0.572

図中の他のアイコンは取り込まれた情報資源を、図中のエッジは、構造化のための関連性を示す。この提示例においては、全ての活動の推定、構造化が適切に行われている。例えば、利用者が「MBL 原稿草案 1」の制作中に参考資料として「多種情報アクセスログにおける同一作業への関連性抽出手法[5]」や Web ページ「What are extensions?」を閲覧していたため、「MBL 原稿草案 1 を制作した活動」による構造化が行われていることが確認できた。

5.2 創造活動構造化機能に関する予備実験評価

前述した実験における創造活動の構造化がどの程度適切に行われているかを検証した。ここでは、前述した「期間中に制作を行った資料」5 個を、実行した創造活動として推定できているか、それぞれの制作中（活動中）に利活用した情報資源を構造化できているかを検証する。

それぞれの資料制作時に利用した情報資源の数を表 3 に示す。なお、本実験では、活動中に利用した全ての種類の情報資源を AIR 化することができなかつたため、検証の対象を画像、文書、スライドのみに絞り、それ以外の情報資源（Web ページ、予定など）は、今後の実験において導入を予定している。対象とした情報資源は全てシステムに取り込まれ、AIR 化された状態で作業を行った。

5 つの活動について、実験の結果より求めた適合率、再現率、F 値の平均を表 4 に示す。表にあるように、適合率は 0.833 となり、誤った構造化がされることは少なかったが、一方で、再現率は 0.436 と低い結果になった。特に、文書の制作時に、図として挿入した画像の構造化がほとんどできていなかった。これは、画像の挿入時にファイルを開くことなく、サムネイルの確認のみで挿入する画像の選択をしていたことが原因だと考えられる。以上より、現状の創造活動構造化機能では再現率が低くなってしまふ場合があり、対処するための改善が必要であることが分かった。

6. おわりに

本稿では、創造活動における情報資源の利活用を支援するための情報資源の構造化獲得支援機構について述べた。また、提案に基づく試作システムの実装を行い、動作の確認および機能の評価を行った。今後は、予備実験で判明し

た問題点に対処するため、構造化機能のアルゴリズムを改善した上で、活動状況に応じた情報資源の構造化の正確性について関連研究と比較・評価を行う。その後、構築された関連性・構造化を活かして、利用者の現在の活動状況に応じた情報提示を行う機能を検討する予定である。

参考文献

- [1] Jin L., Feng L., Liu G. and Wang C., “Personal Web Revisitation by Context and Content Keywords with Relevance Feedback,” IEEE Transactions on Knowledge and Data Engineering, Vol. 29, No. 7, pp. 1508-1521, 2017.
- [2] Deng T., Zhao L., Feng L., “Enhancing Web Revisitation by Contextual Keywords,” International Conference on Web Engineering, ICWE 2013, Lecture Notes in Computer Science, Vol. 7977, pp. 323-337, 2013.
- [3] Halipern, J., Jitkoff, N., Warr, A., Karahalios, K., Sesek, R. and Shkrob, N., “YouPivot: improving recall with contextual search,” Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI’ 11), pp. 1521-1530, 2011.
- [4] 定免睦昌, 國島丈生, 横田一正, “ユーザタスク情報に基づくファイル関連管理手法,” 情報処理学会研究報告, No.1, pp.1-7, 2012.
- [5] 中村明順, 島谷 宙伸, 西尾 信彦, “多種情報アクセスログにおける同一作業への関連性抽出手法,” 電子情報通信学会論文誌 D, Vol. J96-D, No.5, pp.1188-1198, 2013.
- [6] 土田貴裕, 大平茂輝, 長尾確, “ゼミコンテンツの再利用に基づく研究活動支援,” 情報処理学会論文誌, Vol. 51, No. 6, pp. 1357-1370, 2010.
- [7] 石戸田顕太郎, 大平茂輝, 長尾確, “継続的ミーティング支援システム,” 情報処理学会論文誌, Vol. 53, No. 8, pp. 2044-2048, 2012.
- [8] 相原建朗, 堀浩一, “記憶の想起に基づく創造性支援,” 情報処理学会論文誌, Vol. 42, No. 6, pp. 1377-1386, 2001.
- [9] 木下哲男, “分散情報資源活用の一手法-能動的情報資源の設計-,” 信学技報, AI99-54, pp. 13-19, 1999.
- [10] Google Drive API, <https://developers.google.com/drive/> (accessed 2017-10-20).
- [11] Google Calendar API, <https://developers.google.com/google-apps/calendar/> (accessed 2017-10-20).
- [12] Gmail API, <https://developers.google.com/gmail/api/> (accessed 2017-10-20).
- [13] Apache Tika, <https://tika.apache.org/> (accessed 2017-10-20).
- [14] lucene-gosen, <https://github.com/lucene-gosen> (accessed 2017-10-20).
- [15] MongoDB, <https://www.mongodb.com/> (accessed 2017-10-20).
- [16] Chrome extension, <https://developer.chrome.com/extensions> (accessed 2017-10-20).
- [17] Windows API, <https://msdn.microsoft.com/ja-jp/library/cc376921.aspx> (accessed 2017-10-20).